

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 31, 2003, 09:11:17 ; Search time 41 Seconds
(without alignments)
163.052 Million cell updates/sec

Title: US-09-525-041-2
Perfect score: 878
Sequence: 1 MASRMRLLLLLSCLAKTV.....LTWSSNECNKQHFLCKYRP 158

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS-COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	878	100.0	158	2	US-08-729-103-1
2	878	100.0	158	2	US-08-468-413-2
3	878	100.0	158	3	US-09-162-508-2
4	878	100.0	158	5	PCT-US95-07169-2
5	254.5	29.0	165	2	US-08-401-530A-7
6	254.5	29.0	165	2	US-08-729-103-3
7	254.5	29.0	165	2	US-08-709-662-7
8	240	27.3	166	2	US-08-729-103-4
9	231	26.3	174	2	US-08-401-530A-5
10	231	26.3	174	2	US-08-709-662-5
11	228.5	26.0	174	2	US-08-401-530A-6
12	228.5	26.0	174	2	US-08-709-662-6
13	223	25.4	117	6	5514582-15
14	221	25.2	175	2	US-08-464-637-2
15	221	25.2	175	2	US-08-401-530A-4
16	221	25.2	175	2	US-08-709-662-4
17	221	25.2	175	2	US-08-822-261-3
18	221	25.2	175	4	US-09-226-852-3
19	218	24.8	174	1	US-07-778-156-7
20	218	24.8	174	1	US-08-822-261-4
21	218	24.8	174	4	US-08-422-166-7
22	218	24.8	174	4	US-09-226-852-4
23	208	23.7	175	2	US-08-401-530A-3
24	208	23.7	175	2	US-08-709-662-3
25	204.5	23.3	130	1	US-07-893-929A-7
26	204.5	23.3	130	5	PCT-US92-10344-7
27	204	23.2	175	2	US-08-822-261-1

28	204	23.2	175	4	US-09-226-852-1	Sequence 1, Appli
29	198.5	22.6	132	1	US-07-893-929A-5	Sequence 5, Appli
30	198.5	22.6	132	5	PCT-US92-10344-5	Sequence 5, Appli
31	198.5	22.6	134	1	US-07-893-929A-2	Sequence 2, Appli
32	198.5	22.6	134	5	PCT-US92-10344-2	Sequence 2, Appli
33	196	22.3	125	1	US-07-893-929A-3	Sequence 3, Appli
34	196	22.3	125	5	PCT-US92-10344-3	Sequence 3, Appli
35	194.5	22.2	131	1	US-07-893-929A-1	Sequence 1, Appli
36	194.5	22.2	131	5	PCT-US92-10344-1	Sequence 1, Appli
37	193	22.0	912	5	PCT-US95-03747-2	Sequence 1, Appli
38	190	21.6	174	2	US-08-401-530A-2	Sequence 2, Appli
39	190	21.6	174	2	US-08-709-662-2	Sequence 2, Appli
40	190	21.6	175	1	US-08-909-725-6	Sequence 6, Appli
41	189	21.5	123	1	US-07-893-929A-10	Sequence 10, Appli
42	189	21.5	123	5	PCT-US92-10344-10	Sequence 10, Appli
43	189	21.5	273	3	US-09-111-470-10	Sequence 10, Appli
44	189	21.5	292	2	US-08-688-342-4	Sequence 4, Appli
45	189	21.5	292	2	US-09-113-788-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-729-103-1
; Sequence 1, Application US/08729103
; Patent No. 5837841
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN REG PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,103
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0138 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: COLNFET02
; CLONE: 1310334
US-08-729-103-1

Query Match 100.0%; Score 878; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 2e-88;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
Db 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYORSOPITWGLHDPKROQOWIDGAMVLYRSWSG 120
Db 61 YNGAHLASILSLKEASTIAEYISGYORSOPITWGLHDPKROQOWIDGAMVLYRSWSG 120
QY 121 KSMGKNKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158
Db 121 KSMGKNKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158

RESULT 2

US-08-468-413-2
; Sequence 2, Application US/08468413
; Patent No. 5861494
; GENERAL INFORMATION:
; APPLICANT: LI, ET AL.
; TITLE OF INVENTION: Human Colon Specific Gene
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA: US/08/468,413
; APPLICATION NUMBER: 08/468,413
; FILING DATE: 06 JUN 95
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-447
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-468-413-2

Query Match 100.0%; Score 878; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 2e-88;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
Db 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYORSOPITWGLHDPKROQOWIDGAMVLYRSWSG 120
Db 61 YNGAHLASILSLKEASTIAEYISGYORSOPITWGLHDPKROQOWIDGAMVLYRSWSG 120
QY 121 KSMGKNKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158
Db 121 KSMGKNKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158

RESULT 3

US-09-162-508-2
; Sequence 2, Application US/09162508
; Patent No. 6080722
; GENERAL INFORMATION:
; APPLICANT: LI, ET AL.
; TITLE OF INVENTION: Human Colon Specific Gene
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA: US/09/162,508
; APPLICATION NUMBER: US/09/162,508
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/468,413
; FILING DATE: 06 JUN 95
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-447
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-09-162-508-2

Query Match 100.0%; Score 878; DB 3; Length 158;
Best Local Similarity 100.0%; Pred. No. 2e-88;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
Db 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDALEECQS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYORSOPITWGLHDPKROQOWIDGAMVLYRSWSG 120
Db 61 YNGAHLASILSLKEASTIAEYISGYORSOPITWGLHDPKROQOWIDGAMVLYRSWSG 120
QY 121 KSMGKNKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158
Db 121 KSMGKNKHCAMSSNNFLTWSSNECNKROHFLCKYRP 158

RESULT 4

PCT-US95-07169-2
; Sequence 2, Application PC/TUS9507169
; GENERAL INFORMATION:
; APPLICANT: LI, ET AL.
; TITLE OF INVENTION: Human Colon Specific Gene
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD

CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07169
FILING DATE: 06 JUN 95
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-389
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
PCT-1595-07169-2

Query Match 100.0%; Score 878; DB 5; Length 158;

Best Local Similarity 100.0%; Pred. No. 2e-88;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLLSCLAKTGVLDIIMRPSCAPGWYFKSNVCYGYPRKLRNNSDAELECOQ 60
DB 1 MASRSMRLLLLLSCLAKTGVLDIIMRPSCAPGWYFKSNVCYGYPRKLRNNSDAELECOQ 60
QY 61 YNGGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKROQOWIDGAMLYRSWSG 120
DB 61 YNGGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKROQOWIDGAMLYRSWSG 120
QY 121 KSMGKHKCAEMSSNNFLTWSSNCKRQHFCLKYRP 158
DB 121 KSMGKHKCAEMSSNNFLTWSSNCKRQHFCLKYRP 158

RESULT 5
US-08-401-530A-7
Sequence 7, Application US/08401530A
Patent No. 5834590
GENERAL INFORMATION:
APPLICANT: Vinik, Aaron I.
APPLICANT: Pittenger, Gary L.
APPLICANT: Rafaeloff, Ronit
APPLICANT: Rosenberg, Lawrence
APPLICANT: Duguid, William P.
TITLE OF INVENTION: IN GAP PROTEIN INVOLVED IN PANCREATIC
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/401.530A
FILING DATE: 22-FEB-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 00570.48743
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Rattus rattus
US-08-401-530A-7

Query Match 29.0%; Score 254.5; DB 2; Length 165;

Best Local Similarity 30.6%; Pred. No. 3.5e-20;
Matches 52; Conservative 38; Mismatches 51; Indels 29; Gaps 7;

QY 7 RLLLLSCLAKTGVLDIIMRPS-----CAPGWYFKSNVCYGYPRKLRNW 51
DB 5 KYFILLSC-----WVLSPSQGEAEEDLPARITCPGSHNAYSYCYFPMEDHLSW 56
QY 52 SDABLECOYNGGAHLASILSLKEASTIAEYI--SGYORSQPIWIGLHDPKROQOWID 109
DB 57 AEADLFCONMSG-YLVSVLSQAEGNFLASLIKESGTTAAN-VWIGLHDPKRRRHHWS 114
QY 110 GAMVLYRSW-SGKSMGNK-HCAEMSSNNFLTWSSNCKRQHFCLKYR 157
DB 115 GSLFLYKSWDTGYPPNNSNRGYCVTSNSGYKWRDSCDAQLGFVCKFK 164

RESULT 6
US-08-729-103-3
Sequence 3, Application US/08729103
Patent No. 5837841
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN REG PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,103
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0138 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

```
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 393209
US-08-729-103-3

Query Match          29.0%; Score 254.5; DB 2; Length 165;
Best Local Similarity 30.6%; Pred. No. 3.5e-20;
Matches 52; Conservative 38; Mismatches 51; Indels 29; Gaps 7;

QY 7 RLLLSLCLAKTGVLDIIMRPS-----CAPGWFYHKSNCYGFRLRNW 51
Db 5 KYFILLSCL-----MVLSPSQGQAEEDLPARITCPEGSNAYSSCYTFMEDHLSW 56
QY 52 SDAELECSYGNCAHLASILSLKEASTIAEYI--SGYORSQPIWIGLHDPQKQOWID 109
Db 57 AADLFQNNNSG-YLVSVLSQAEGNFLASLIKESGTTAAN-VWIGLHDPKNNRRHWSS 114
QY 110 GAMLYRWSW-SGKSMGNGK-HCAEMSSNNNFLTWSNECNKQKQHFCKYR 157
Db 115 GSLFLYKSWDTGYPNNSNRGVCVSTNSGYKKWRDNSCDAQLSFVCKPK 164

RESULT 7
US-08-709-662-7
; Sequence 7, Application US/08709662
; Patent No. 5840531
; GENERAL INFORMATION:
; APPLICANT: Vinik, Aaron I.
; APPLICANT: Pittenger, Gary L.
; APPLICANT: Rafaeloff, Ronit
; APPLICANT: Rosenberg, Lawrence
; APPLICANT: Duguid, William P.
; TITLE OF INVENTION: INSP PROTEIN INVOLVED IN PANCREATIC
; TITLE OF INVENTION: ISLET NEOGENESIS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,662
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 00570.59178
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Rattus rattus

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 393209
US-08-729-103-3

Query Match          29.0%; Score 254.5; DB 2; Length 165;
Best Local Similarity 30.6%; Pred. No. 3.5e-20;
Matches 52; Conservative 38; Mismatches 51; Indels 29; Gaps 7;

QY 7 RLLLSLCLAKTGVLDIIMRPS-----CAPGWFYHKSNCYGFRLRNW 51
Db 5 KYFILLSCL-----MVLSPSQGQAEEDLPARITCPEGSNAYSSCYTFMEDHLSW 56
QY 52 SDAELECSYGNCAHLASILSLKEASTIAEYI--SGYORSQPIWIGLHDPQKQOWID 109
Db 57 AADLFQNNNSG-YLVSVLSQAEGNFLASLIKESGTTAAN-VWIGLHDPKNNRRHWSS 114
QY 110 GAMLYRWSW-SGKSMGNGK-HCAEMSSNNNFLTWSNECNKQKQHFCKYR 157
Db 115 GSLFLYKSWDTGYPNNSNRGVCVSTNSGYKKWRDNSCDAQLSFVCKPK 164

RESULT 8
US-08-729-103-4
; Sequence 4, Application US/08729103
; Patent No. 5837841
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN REG PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,103
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0138 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 474306
US-08-729-103-4

Query Match          27.3%; Score 240; DB 2; Length 166;
Best Local Similarity 34.3%; Pred. No. 1.3e-18;
Matches 46; Conservative 27; Mismatches 57; Indels 4; Gaps 4;

QY 27 RPSCAPGWFYHKSNCYGFRLRNWSDAELECSYGNCAHLASILSLKEASTIAEYISGY 86
Db 33 RISCPEGTNAYRSYCYTFMEDPETWVDADLYQNNMNSG-NLVSVLTAQEGAFVASLIKES 91
```

QY 87 QRSQP-IWIGLHDPKQKQWQWIDGAMVLYRSW-SGKSMGNGK-HCAEMSSNNFLTWSS 143
Db 92 STDDSNNVWIGLHDPKQKQWQWIDGAMVLYRSW-SGKSMGNGK-HCAEMSSNNFLTWSS 151
QY 144 NECNKRQHFLCKYR 157
Db 152 ESCEKKFVCKFK 165

RESULT 9
US-08-401-530A-5
; Sequence 5, Application US/08401530A
; Patent No. 5834590
; GENERAL INFORMATION:
; APPLICANT: Vinik, Aaron I.
; APPLICANT: Pittenger, Gary L.
; APPLICANT: Rafaeloff, Ronit
; APPLICANT: Rosenberg, Lawrence
; APPLICANT: Duguid, William P.
; TITLE OF INVENTION: IN GAP PROTEIN INVOLVED IN PANCREATIC
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/401,530A
; FILING DATE: 22-FEB-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 00570.48743
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Rattus rattus

US-08-401-530A-5
Query Match 26.3%; Score 231; DB 2; Length 174;
Best Local Similarity 32.4%; Pred. No. 1.4e-17;
Matches 55; Conservative 28; Mismatches 73; Indels 14; Gaps 5;
QY 1 MASRSMRLLLLSCLAKTGVLD-----IMRPSCAPGWFFYHKSNVCYGFRLRNWSDA 54
Db 5 VALTTMSWMLLSLMLLSQVQGEDAKEDVPTSRISCPKGSRAYGSYCYALFVSVKSFWDA 64
QY 55 ELECSYNGAHLASILSLKEASTIAEYI-SGYQRSQPIWIGLHDPKQKQ-----WQWI 108
Db 65 DLACQKRPSSG-HLVSVLSGSEASFVSSLLKSGNSGQNVWIGLHDPKQKQ-----WQWI 123
QY 109 DGAMVLYRSW-SGKSMGNGKHCAEMSSNNFLTWSSNECNKQHFCLCKYR 157
Db 124 NADVNYFNWETNPSSVSGSHCGTLTRASGFLRWRENNCISELPYVCKFK 173

RESULT 10
US-08-401-530A-6
; Sequence 6, Application US/08401530A
; Patent No. 5834590
; GENERAL INFORMATION:
; APPLICANT: Vinik, Aaron I.
; APPLICANT: Pittenger, Gary L.
; APPLICANT: Rafaeloff, Ronit
; APPLICANT: Rosenberg, Lawrence
; APPLICANT: Duguid, William P.
; TITLE OF INVENTION: IN GAP PROTEIN INVOLVED IN PANCREATIC

US-08-709-662-5
; Sequence 5, Application US/08709662
; Patent No. 5840531
; GENERAL INFORMATION:
; APPLICANT: Vinik, Aaron I.
; APPLICANT: Pittenger, Gary L.
; APPLICANT: Rafaeloff, Ronit
; APPLICANT: Rosenberg, Lawrence
; APPLICANT: Duguid, William P.
; TITLE OF INVENTION: IN GAP PROTEIN INVOLVED IN PANCREATIC
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,662
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 00570.59178
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Rattus rattus

US-08-709-662-5
Query Match 26.3%; Score 231; DB 2; Length 174;
Best Local Similarity 32.4%; Pred. No. 1.4e-17;
Matches 55; Conservative 28; Mismatches 73; Indels 14; Gaps 5;
QY 1 MASRSMRLLLLSCLAKTGVLD-----IMRPSCAPGWFFYHKSNVCYGFRLRNWSDA 54
Db 5 VALTTMSWMLLSLMLLSQVQGEDAKEDVPTSRISCPKGSRAYGSYCYALFVSVKSFWDA 64
QY 55 ELECSYNGAHLASILSLKEASTIAEYI-SGYQRSQPIWIGLHDPKQKQ-----WQWI 108
Db 65 DLACQKRPSSG-HLVSVLSGSEASFVSSLLKSGNSGQNVWIGLHDPKQKQ-----WQWI 123
QY 109 DGAMVLYRSW-SGKSMGNGKHCAEMSSNNFLTWSSNECNKQHFCLCKYR 157
Db 124 NADVNYFNWETNPSSVSGSHCGTLTRASGFLRWRENNCISELPYVCKFK 173

RESULT 11
US-08-401-530A-6
; Sequence 6, Application US/08401530A
; Patent No. 5834590
; GENERAL INFORMATION:
; APPLICANT: Vinik, Aaron I.
; APPLICANT: Pittenger, Gary L.
; APPLICANT: Rafaeloff, Ronit
; APPLICANT: Rosenberg, Lawrence
; APPLICANT: Duguid, William P.
; TITLE OF INVENTION: IN GAP PROTEIN INVOLVED IN PANCREATIC

;; TITLE OF INVENTION: ISLET NEOGENESIS
;; NUMBER OF SEQUENCES: 7
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Banner & Allegretti
;; STREET: 1001 G Street, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: US
;; ZIP: 20001-4597
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/401.530A
;; FILING DATE: 22-FEB-1995
;; CLASSIFICATION: 800
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kagan, Sarah A.
;; REGISTRATION NUMBER: 32,141
;; REFERENCE/DOCKET NUMBER: 00570.48743
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-508-9100
;; TELEFAX: 202-508-9299
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 174 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; ORIGINAL SOURCE:
;; ORGANISM: Rattus rattus
;; US-08-401-530A-6

Query Match 26.0%; Score 228.5; DB 2; Length 174;
Best Local Similarity 34.5%; Pred. No. 2.6e-17;
Matches 48; Conservative 25; Mismatches 57; Indels 9; Gaps 4;

QY 27 RPSCAPGWFYHKSNVCYGRKLRNWSDALEQCSYGNAGHLASLSLKEASTIAEYISG- 85
DB 36 RTSCPMGSKAYRSYCYTLVTLKSWFQADLACQKRPBG-HLVSLISGGEASFVSSLVGTGR 94
QY 86 YORSQPIWIGLHDPQKRQ-----WOWIDGAMVLYRSWSG--KSMGKNKHCAMSSNNNF 138
DB 95 VNNNQDIWIWLHPTMGQQPNGGCGWNSDVLNLYLNWDGDPSSTVNRGCGSLTATSEF 154
QY 139 LTWSSNECNKQHFLLCKYR 157
DB 155 LKMGDHHCDVELPFCVKFK 173

RESULT 12

US-08-709-662-6
Sequence 6, Application US/08709662
Patent No. 5840531

;; GENERAL INFORMATION:
;; APPLICANT: Vinik, Aaron I.
;; APPLICANT: Pittenger, Gary L.
;; APPLICANT: Rafaeloff, Ronit
;; APPLICANT: Rosenberg, Lawrence
;; APPLICANT: Duguid, William P.
;; TITLE OF INVENTION: IN GAP PROTEIN INVOLVED IN PANCREATIC
;; TITLE OF INVENTION: ISLET NEOGENESIS
;; NUMBER OF SEQUENCES: 7
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Banner & Witcoff, Ltd.
;; STREET: 1001 G Street, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: US
;; ZIP: 20001-4597
;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/709,662
;; FILING DATE: 09-SEP-1996
;; CLASSIFICATION: 435.
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kagan, Sarah A.
;; REGISTRATION NUMBER: 32,141
;; REFERENCE/DOCKET NUMBER: 00570.59178
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-508-9100
;; TELEFAX: 202-508-9299
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 174 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; ORIGINAL SOURCE:
;; ORGANISM: Rattus rattus
;; US-08-709-662-6

Query Match 26.0%; Score 228.5; DB 2; Length 174;
Best Local Similarity 34.5%; Pred. No. 2.6e-17;
Matches 48; Conservative 25; Mismatches 57; Indels 9; Gaps 4;

QY 27 RPSCAPGWFYHKSNVCYGRKLRNWSDALEQCSYGNAGHLASLSLKEASTIAEYISG- 85
DB 36 RTSCPMGSKAYRSYCYTLVTLKSWFQADLACQKRPBG-HLVSLISGGEASFVSSLVGTGR 94
QY 86 YORSQPIWIGLHDPQKRQ-----WOWIDGAMVLYRSWSG--KSMGKNKHCAMSSNNNF 138
DB 95 VNNNQDIWIWLHPTMGQQPNGGCGWNSDVLNLYLNWDGDPSSTVNRGCGSLTATSEF 154
QY 139 LTWSSNECNKQHFLLCKYR 157
DB 155 LKMGDHHCDVELPFCVKFK 173

RESULT 13

5514582-15
Patent No. 5514582

;; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
;; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
;; IMMUNOGLOBULINS

;; NUMBER OF SEQUENCES: 43
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/185,670
;; FILING DATE: 21-JAN-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 986,931
;; FILING DATE: 08-DEC-1992
;; APPLICATION NUMBER: 808,122
;; FILING DATE: 16-DEC-1991
;; APPLICATION NUMBER: 440,625
;; FILING DATE: 22-NOV-1989
;; APPLICATION NUMBER: 315,015
;; FILING DATE: 23-FEB-1989
;; SEQ ID NO:15;
;; LENGTH: 117

5514582-15

Query Match 25.4%; Score 223; DB 6; Length 117;
Best Local Similarity 34.7%; Pred. No. 6.1e-17;
Matches 41; Conservative 30; Mismatches 41; Indels 6; Gaps 5;

QY 41 CYGFRKLRNWSDALEQCSYGNAGHLASLSLKEASTIAEYI--SGYORSQPIWIGLHD 98
DB 2 CYFMEDHLSWAEDLFCQNMNSG-YLVSVLSWAEGNFLASLKESGTTAAN-VWIGLHD 59

QY 99 PQKRWQWIDGAMYLRSW-SGKSMGNGK-HCAEMSSNNFLTWSSNECNKRQHPLC 154
 Db 60 PKNRRWHSSGSLFLYKSWDTGTPNNRNGYCVTSNYSYKWKRDNSCDAQLSFVC 117

RESULT 14

US-08-464-637-2
 ; Sequence 2, Application US/08464637
 ; Patent No. 5834214
 ; GENERAL INFORMATION:
 ; APPLICANT: Iovanna, Juan-Lucio
 ; APPLICANT: Dagorn, Jean-Charles
 ; APPLICANT: Keim, Volker
 ; APPLICANT: Sarles, Jacques
 ; TITLE OF INVENTION: Detection of Pancreatitis-Associated
 ; TITLE OF INVENTION: Protein for Diagnosis of Cystic Fibrosis or Pancreatic
 ; TITLE OF INVENTION: Disease (as amended).
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
 ; STREET: P.O. Box 747
 ; CITY: Falls Church
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22040-3487

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/464,637
 ; FILING DATE: 30-AUG-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murphy Jr., Gerald M.
 ; REGISTRATION NUMBER: 28,977
 ; REFERENCE/DOCKET NUMBER: 2121-107P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-205-8000
 ; TELEFAX: 703-205-8050
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 175 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; ORIGIN: GenBank
 ; ORGANISM: Homo sapiens
 ; US-08-464-637-2

Query Match 25.2%; Score 221; DB 2; Length 175;
 Best Local Similarity 32.7%; Pred. No. 1.7e-16;
 Matches 53; Conservative 29; Mismatches 64; Indels 16; Gaps 7;

QY 10 LLLSCL-AKTGVLDIIMR--PS-----CAPGWFYHKSNCYGYFRKLWNWSDAELECSYSG 62
 Db 13 MLLSCLMLLSQVQGEPPQELPSARIRCPKSKAYGSHCYALFLSPKSWTDADLACQKRP 72
 QY 63 NGAHLASILSLKEASTIAYISGYVSQP-IWGLHDPKQKQ-----WQWIDGAMYLRY 116
 Db 73 SG-NLVSVLGAGSGFVSSLVKSIGNSYVWIGLHDPQTQGTPEPNEGGEWSSSDVMNYF 131
 QY 117 SW--SGKSMGNGKHCAMSSNNFLTWSSNECNKRQHFLCKY 156
 Db 132 AMERNPSTISSPGHCASLRSSTAFLRWKDYNVNCNRLPYVCKF 173

RESULT 15

US-08-401-530A-4
 ; Sequence 4, Application US/08401530A
 ; Patent No. 5834590
 ; GENERAL INFORMATION:
 ; APPLICANT: Vinik, Aaron I.
 ; APPLICANT: Pittenger, Gary L.

APPLICANT: Rafaeloff, Ronit
 APPLICANT: Rosenberg, Lawrence
 APPLICANT: Duguid, William P.
 TITLE OF INVENTION: INCAP PROTEIN INVOLVED IN PANCREATIC
 TITLE OF INVENTION: ISLET NEOGENESIS
 NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Banner & Allegretti
 STREET: 1001 G Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: US
 ZIP: 20001-4597

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/401,530A
 ; FILING DATE: 22-FEB-1995
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kagan, Sarah A.
 ; REGISTRATION NUMBER: 32,141
 ; REFERENCE/DOCKET NUMBER: 00570.48743
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-508-9100
 ; TELEFAX: 202-508-9299
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 175 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; US-08-401-530A-4

Query Match 25.2%; Score 221; DB 2; Length 175;
 Best Local Similarity 32.7%; Pred. No. 1.7e-16;
 Matches 53; Conservative 29; Mismatches 64; Indels 16; Gaps 7;

QY 10 LLLSCL-AKTGVLDIIMR--PS-----CAPGWFYHKSNCYGYFRKLWNWSDAELECSYSG 62
 Db 13 MLLSCLMLLSQVQGEPPQELPSARIRCPKSKAYGSHCYALFLSPKSWTDADLACQKRP 72
 QY 63 NGAHLASILSLKEASTIAYISGYVSQP-IWGLHDPKQKQ-----WQWIDGAMYLRY 116
 Db 73 SG-NLVSVLGAGSGFVSSLVKSIGNSYVWIGLHDPQTQGTPEPNEGGEWSSSDVMNYF 131
 QY 117 SW--SGKSMGNGKHCAMSSNNFLTWSSNECNKRQHFLCKY 156
 Db 132 AMERNPSTISSPGHCASLRSSTAFLRWKDYNVNCNRLPYVCKF 173

Search completed: December 31, 2003, 09:16:48
 Job time : 43 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 31, 2003, 09:22:09 ; Search time 2056 Seconds
(without alignments)
1867.756 Million cell updates/sec

Title: US-09-525-041-2
Perfect score: 878
Sequence: 1 MASRSNRLLLLSCLAKTV.....LTWSSNECNKROHFLCKYRP 158

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=xlp
-O=/cgn2_1/USPTO.spool_p/US09525041/runat_31122003_091149_11877/app_query.fasta_1.327
-DB=EST -QPMR=fastap -SUFFIX=rest -MINMATCH=0.1 -LOPCU=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09525041 @CGN 1.1 3549 @runat_31122003_091149_11877 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DSEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_esti:*
10: gb_estc:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	878	100.0	663	12	BM764955	BM764955 K-EST0046
2	878	100.0	708	12	BM772617	BM772617 K-EST0036
3	878	100.0	822	12	BG770898	BG770898 602719106
4	873	99.4	665	12	BM764658	BM764658 K-EST0046
5	869	99.0	628	12	BM739168	BM739168 K-EST0008
6	859	97.8	600	12	BM852931	BM852931 K-EST0134
7	858	97.7	667	12	BM854067	BM854067 K-EST0136
8	852	97.0	835	10	BE872152	BE872152 601446259
9	834	95.0	664	10	AW957953	AW957953 EST370023
10	822	93.6	908	10	BG253845	BG253845 602366542
11	792	90.2	580	12	BM788204	BM788204 K-EST0067
12	785	89.4	551	9	AA314779	AA314779 EST186601
13	785	89.4	968	10	BG256113	BG256113 602367238
14	761	86.7	440	9	AA581222	AA581222 nd38609.r
15	752	85.6	526	9	AI791498	AI791498 ni2310.Y
16	752	85.6	527	9	AI791286	AI791286 ae62509.Y
17	752	85.6	557	12	BM769915	BM769915 K-EST0053
18	734	83.6	535	12	BM852851	BM852851 K-EST0134
19	730	83.1	546	12	BM821464	BM821464 K-EST0090
20	717	81.7	532	9	AA308533	AA308533 EST179533
21	716	81.5	508	9	AA315990	AA315990 EST187724
22	716	81.5	511	9	AA314573	AA314573 EST186473
23	685	78.0	514	9	AA308019	AA308019 EST178882
24	684	77.9	502	9	AA307789	AA307789 EST178855
25	678	77.2	517	9	AA315862	AA315862 EST187607
26	663	75.5	512	12	BM788199	BM788199 K-EST0067
27	657	74.8	356	9	AA315757	AA315757 EST187733
28	652	74.3	454	9	AA316634	AA316634 EST188314
29	652	74.3	459	9	AA313549	AA313549 EST185429
30	652	74.3	503	12	BM769144	BM769144 K-EST0052
31	648	73.8	483	9	AA316848	AA316848 EST188510
32	648	73.8	488	12	BM792702	BM792702 K-EST0073
33	648	73.8	492	12	AA313200	AA313200 EST185134
34	645	73.5	489	12	BM796494	BM796494 K-EST0079
35	628	71.5	460	9	AW363214	AW363214 CM3-CT031
36	615	70.0	370	14	US4601	US4601 HSU54601 HU
37	614	69.9	600	14	CB288128	CB288128 CMD74_E10
38	613	69.8	482	9	AA316249	AA316249 EST188027
39	600.5	68.4	702	14	BY708034	BY708034 BY708034
40	600.5	68.4	1013	11	AK008049	AK008049 Mus muscu
41	598.5	68.2	535	10	BE138108	BE138108 ug48f05.Y
42	594.5	67.7	1024	11	AK008438	AK008438 Mus muscu
43	592	67.4	694	12	BM790284	BM790284 K-EST0070
44	591.5	67.4	545	4	BX513643	Bx513643 RZPD Mus
45	589.5	67.1	741	14	BY708339	BY708339 BY708339

ALIGNMENTS

RESULT 1
BM764955
LOCUS
DEFINITION K-EST0046563 S3SNU1661 Homo sapiens CDNA clone S3SNU1661-12-E07 5',
663 bp mRNA linear EST 04-MAR-2002
rna sequence.
ACCESSION BM764955
VERSION BM764955.1 GI:19094570
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 663)

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4400
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 12 row: E column: 07
High quality sequence stop: 663.

FEATURES

source

1. .663
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S3SNU16s1-12-E07"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/cell_line="SNU-16"
/lab_host="Top10F"
/clone_lib="S3SNU16s1"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 ~ 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(dT)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the subtraced cDNA
libraries were constructed by transformation of the
remaining DNA into competent cells E. coli Top10F, with
electroporation method."

BASE COUNT 182 a 159 c 180 g 142 t
ORIGIN

Alignment Scores:
Pred. No.: 5.65e-90 Length: 663
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-525-041-2 (1-158) x BM764955 (1-663)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
|||||
Db 147 ATGGCTTCCAGAGCATGCGGCTGCTCTATGTGAGCTGCCGCGCAAAACAGGAGTC 206
|||||
Qy 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTTPheTyrHisLysSerAsn 40
|||||

Db 207 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTCGATGGTTTTTACCACAAGTCCAAT 266
|||||
Qy 41 CyeTyrGlyTyrPheArgLysLeuArgAsnTTPSerAspAlaGluLeuGluCysGlnSer 60
|||||
Db 267 TGTATGGTTTACTTCAGGAAGCTGAGAACTGGTCTGTATGCCAGGCTCGAGTGTCACTCT 326
|||||
Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
|||||
Db 327 TAGGGAACGGAGCCACCTGGCATCTATCTCTGAGTTTAAAGGAAGCCAGCACCACATAGCA 386
|||||
Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTTPileGlyLeuHisAppProGln 100
|||||
Db 387 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCTCGACGACCCACAG 446
|||||
Qy 101 LysArgGlnGlnTTPGlnTTPileAspGlyAlaMetTyrLeuTyrArgSerTTPSerGly 120
|||||
Db 447 AAGAGGAGCAGTGGGAGTGGATTGTATGGGGCCATGTATCTGTACAGATCCTGGTCTGGC 506
|||||
Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
|||||
Db 507 AAGTCCATGGTGGGAAACAGCACTGTGCTGAGATGAGCTCCATTAACAACCTTTTAACT 566
|||||
Qy 141 TTPSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
|||||
Db 567 TGGAGCAGCAACGATGCAACAGCGCAACACTTCTGTGCAAGTACCGACCA 620
|||||

RESULT 2

BM772617

LOCUS

DEFINITION

K-EST0056822 S3SNU16s1 Homo sapiens cDNA clone S3SNU16s1-14-G04 5',

mRNA sequence.

ACCESSION

BM772617

VERSION

BM772617.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 708)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and

Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 14 row: G column: 04

High quality sequence stop: 708.

Location/Qualifiers

1..708

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S3SNU16s1-14-G04"

/sex="F"

/tissue_type="Ascites"

/cell_type="Lymphoblast-like"

/cell_line="SNU-16"

/lab_host="Top10F"

/clone_lib="S3SNU16s1"

/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;

Site 2: NotI; The poly (A)+ RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP) and then decapped

with tabacco acid pyrophosphatase (TAP). The decapped

intact mRNA was ligated with DNA-RNA linker including EcoR

I site by treatment of T4 RNA ligase and the first strand

cDNA was synthesized from oligo dT-selected mRNA by

priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10⁺ by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(drf)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10⁺ with electroporation method."

BASE COUNT 190 a 174 c 185 g 159 t
ORIGIN

Alignment Scores:
Pred. No.: 6,23e-90 Length: 708
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-525-041-2 (1-158) x BM772617 (1-708)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB 149 ATGGCTTCAGAGAGCATGCGGCTGCTTATTCGAGTGCCTGCCAAACAGGAGTC 208
QY 21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTrpHisLysSerAsn 40
DB 209 CTGGGTGATATCATATGAGACCCAGCTGTGCTCTGGATGGTGTATACCAAGTCCAAT 268
QY 41 CysTrpGlyTrpPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
DB 269 TGCTATGTTTACTTCAGGAAGCTGAGGAACCTGTGCTGATGCCGAGCTCGAGTCTAGTCT 328
QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB 329 TACGGAAACGGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 388
QY 81 GluTrpIleSerGlyTrpGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
DB 389 GAGTACATAGTGGCTATCAGAGAAGCCAGCCCATATGGATTGGCTGCAAGCCACAG 448
QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTrpArgSerTrpSerGly 120
DB 449 AAGAGGCAGCAGTGGCAGTGGATTGATGGGCCATGATCTATCTGACAGATCTGCTGGC 508
QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnLeuPheLeuThr 140
DB 509 AAGTCCATGGTGGGAACAGCACTGTGCTGAGATGAGTCCCAATAACAACATTTTAACT 568
QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTrpArgPro 158
DB 569 TGGAGCAGCAACGAATGCAACAGCGCCAACTTCTCTGTGCAAGTACCGACCA 622

RESULT 3

LOCUS BG770898

DEFINITION 602719106F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4839673 5',

mRNA sequence.

ACCESSION BG770898

VERSION
KEYWORDS
SOURCE
ORGANISM

BG770898.1 GI:14081551

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 822)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM1669 row: k column: 02

High quality sequence stop: 731.

FEATURES

Source

1..822

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4839673"

/tissue_type="adenocarcinoma"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NIH_MGC_60"

/note="Organ: prostate; Vector: pDNR-LIB (Clontech);

Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggccattatggcc

); Double-stranded cDNA was prepared from cell line RNA.

5' and 3' adaptors were used in cloning as follows: 5'

adaptor sequence: 5'-CAGCGCATTTATGCC-3' and 3' adaptor

sequence: 5'-ATTCTAGGCGGAGCGCGGCACATG-dT(30)BN-3'

(where B = A, C, or G and N = A, C, G, or T). Average

insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies

contained inserts by PCR. This library was enriched for

full-length clones and was constructed by Clontech

Laboratories (Palo Alto, CA). Note: this is a NIH_MGC

Library."

BASE COUNT 224 a 197 c 217 g 184 t

ORIGIN

Alignment Scores:

Pred. No.: 7,79e-90 Length: 822

Score: 878.00 Matches: 158

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

US-09-525-041-2 (1-158) x BG770898 (1-822)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB 148 ATGGCTTCAGAGAGCATGCGGCTGCTTATTCGAGTGCCTGCCAAACAGGAGTC 207
QY 21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTrpHisLysSerAsn 40
DB 208 CTGGGTGATATCATATGAGACCCAGCTGTGCTCTGGATGGTGTATACCAAGTCCAAT 267
QY 41 CysTrpGlyTrpPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
DB 268 TGCTATGTTTACTTCAGGAAGCTGAGGAACCTGTGCTGATGCCGAGCTCGAGTCTAGTCT 327
QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB 328 TACGGAAACGGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 387
QY 81 GluTrpIleSerGlyTrpGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100

388 GAGTACATAAGTGGCTATCAGAGAGCCAGCCGATATGGATTGGCTTCACAGCCACAG 447
 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 448 AAGAGGCGAGCAGTGGCAGTATGATGGGCCATGATCTGTACAGATCTGTCTGGC 507
 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPhleLeuThr 140
 508 AAGTCCATGGTGGGAACAAGCAGCTGTCTGAGATGAGTCCATCAATAACAACCTTTTAACT 567
 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuLysTyrArgPro 158
 568 TGGAGCAGCAAGCAATCAACAAGCCCAACACTTCTGTGCAAGTACCGACCA 621

RESULT 4
 BM764658
 LOCUS K-EST0046207 S3SNUI6s1 Homo sapiens cDNA clone S3SNUI6s1-7-G11 5',
 DEFINITION mRNA sequence.
 ACCESSION BM764658
 VERSION BM764658.1 GI:19094273
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 665)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 7 row: G column: 11
 High quality sequence stop: 665.
 Location/Qualifiers
 1. 665
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S3SNUI6s1-7-G11"
 /sex="F"
 /tissue_type="Ascites"
 /cell_type="Lymphoblast-like"
 /lab_host="Top10F"
 /clone_lib="S3SNUI6s1"
 /note="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library. After analyzing and
 sequencing about 2,000 ~ 3,000 colonies in original cDNA
 library, the abundant cDNAs were selected and amplified by
 PCR reaction using vector region primer including T7
 promoter as 5' primer and N(dt)14 as 3' primer. The PCR

products were used as template for synthesis of
 biotinylated single stranded RNA by in vitro transcription
 reaction. The synthesized RNA probes were hybridized with
 antisense single stranded cDNAs prepared from original
 library and incubated with avidin-gel. After removing
 DNA-RNA hybrids by centrifuge, the subtracted cDNA
 libraries were constructed by transformation of the
 remaining DNA into competent cells E. coli Top10F' with
 electroporation method."

BASE COUNT 184 a 160 c 180 g 140 t 1 others
 ORIGIN

Alignment Scores: 2,12e-89 Length: 665
 Pred. No.: 873.00 Matches: 157
 Score: 873.00
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 1
 Query Match: 99.43% Indels: 0
 DB: 12 Gaps: 0

US-09-525-041-2 (1-158) x BM764658 (1-665)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 149 ATGGCTTCCAGAAAGCATGGCGTGTCTTATTGCTGAGCTGCTGCCAAACAGGAGTC 208
 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 209 CTGGGTGATATCATCATGAGACCAGCTGTGCTCTGGATGGTTTTACCAACAGTCCAT 268
 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 269 TGCTATGGTTTACTTCAGGAAGCTGAGAACTGCTGTGATGCCGAGCTCGAGTGTCACTCT 328
 61 TyrGlyAsnGlyAlaHisIleAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 329 TACGGAAACGGAGCCACCTGGCATCTATCTCTGAGTTTAAAGGAAGCCAGCACCATAGCA 388
 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 389 GAGTACATAAGTGGCTATCAGAAAGCCAGCCCATATGATGGCTTCGACGCCACAG 448
 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 449 AAGAGGCGAGCAGTGGCAGTGGATTGATGGGCCATGATCTGTACAGATCTGTCTGGC 508
 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPhleLeuThr 140
 509 AAGTCCATGGTGGGAACAAGCAGCTGTCTGAGATGAGTCCATCAATAACAACCTTTTAACT 568
 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuLysTyrArgPro 158
 569 TGGAGCAGCAAGCAATCAACAAGCCCAACACTTCTGTGCAAGTACCGACCA 622

RESULT 5
 BM739168
 LOCUS K-EST0008559 S3SNUI6 Homo sapiens cDNA clone S3SNUI6-3-G07 5', mRNA
 DEFINITION sequence.
 ACCESSION BM739168
 VERSION BM739168.1 GI:19060497
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 628)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished
 Contact: Kim YS

Qy 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 Db 196 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTCTGGATGTTTACCACAAGTCCAAT 255
 Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 Db 256 TGCTATGTTTACTTTCAGGAAGCTGAGGAACCTGGTCTGATGCCGAGCTCGAGTGTCACTCT 315
 Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 Db 316 TACGGAAACGGAGCCACCTGGCATCTATCTCTGAGTTTAAAGGAAGCAGCACCATTAGCA 375
 Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 Db 376 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATGGCTGCAGCACCACAG 435
 Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 Db 436 AAGAGGAGCAGTGGCAGTGGATTCATGGGGCCATGTATCTGTACAGATCTCTGCTGGC 495
 Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnPheLeuThr 140
 Db 496 AAGTCCATAGTGGTGGGAACAACACTGTGCTGAGATGAGCTCCATATAACAACCTTTTAACT 555
 Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLys 155
 Db 556 TGGAGCAGCAACGATGCACAGCGCCACACTTCTCTGTGCAAG 600

RESULT 7
 BM854067
 LOCUS K-EST0136364 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-145-C05 5', mRNA sequence.
 BM854067
 VERSION BM854067.1 GI:19210466
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 667)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.krribb.re.kr
 Plate: 145 row: C column: 05
 High quality sequence stop: 667.
 Location/Qualifiers
 1..667
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S22SNU16n1-145-C05"
 /sex="F"
 /tissue_type="Ascites"
 /cell_type="Lymphoblast-like"
 /lab_host="SNU-16"
 /clone_lib="S22SNU16n1"
 /note="Organ: Stomach; Vector: pYT3-Pac; Site 1: EcoRI; Site 2: NotI; The S22SNU16 library was contributed by the Soares laboratory and it was constructed as described by Ronaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cells of SNU-16 culture. SNU-16 cell was obtained from

FEATURES

source
 1..667
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S22SNU16n1-145-C05"
 /sex="F"
 /tissue_type="Ascites"
 /cell_type="Lymphoblast-like"
 /lab_host="SNU-16"
 /clone_lib="S22SNU16n1"
 /note="Organ: Stomach; Vector: pYT3-Pac; Site 1: EcoRI; Site 2: NotI; The S22SNU16 library was contributed by the Soares laboratory and it was constructed as described by Ronaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cells of SNU-16 culture. SNU-16 cell was obtained from

Korean Cell Line Bank (KCLB). SNU-16 was established from ascitic fluids of Korean patients by Park J.G. et al. (1990), Cancer Res 50: 2773-2780."

BASE COUNT 180 a 168 c 166 g 153 t
 ORIGIN

Alignment Scores:
 Pred. No.: 11e-87 Length: 667
 Score: 858.00 Matches: 154
 Percent Similarity: 98.10% Conservative: 1
 Best Local Similarity: 97.47% Mismatches: 3
 Query Match: 97.72% Indels: 0
 DB: 12 Gaps: 0

US-09-525-041-2 (1-158) x BM854067 (1-667)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 Db 112 ATGGCTTCCAGAAAGCATGGCGTCTCTATTCCTGAGCTGCCTGGCCAAACAGAGATC 171
 Qy 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 Db 172 CTGGGTGATATCATCATGAGACCCATCTGTGCTCTCTGGATGGTTTACCACAAGTCCAAT 231
 Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 Db 232 TGCTATGTTTACTTCAAGAAGCTGAGAACTGTGCTGATGCCGAGCTCGAGTGTCACTCT 291
 Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 Db 292 TACGGAAACGGAGCCACCTGGCATCTATCTCTGAGTTTAAAGGAAGCAGCACCATTAGCA 351
 Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 Db 352 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCCATATGGATGGCTGCAGCACCACAG 411
 Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 Db 412 AAGACGAGCAGTGGCAGCTGGATTGATGGGCCATGTATCTGTACAGATCTCTGGTCTGGC 471
 Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnPheLeuThr 140
 Db 472 AAGTCCATGTTGGGAAACAGCACTGTGCTGAGATGAGCTCCATATAACAACCTTTTAACT 531
 Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
 Db 532 TGGAGCAGCAACGATGCACAGCGCCACACTTCTCTGTGCAAGTACCACCA 585

RESULT 8
 BE872152
 LOCUS 601446259F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3850358 5', mRNA linear EST 20-OCT-2000
 DEFINITION mRNA sequence.
 BE872152
 ACCESSION BE872152
 VERSION BE872152.1 GI:10320928
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 835)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: L1AM9569 row: e column: 15
 High quality sequence stop: 613.
 Location/Qualifiers
 1. .835
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3850358"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 65"
 /notes="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.8 kb. Library constructed by Life Technologies."

BASE COUNT	226 a	218 c	216 g	175 t
ORIGIN				

 Alignment Scores:
 Pred. No.: 7,49e-87 Length: 835
 Score: 852.00 Matches: 158
 Percent Similarity: 98.75% Conservative: 0
 Best Local Similarity: 98.75% Mismatches: 0
 Query Match: 97.04% Indels: 2
 DB: 10 Gaps: 0

US-09-525-041-2 (1-158) x BE872152 (1-835)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 DB 131 ATGGCTTCCAGAGAGCTGCGGCTGCTCTATATGCTGAGCTGCGCCAAACAGGAGTC 190

QY 21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 DB 191 CTGGGTGATATCATCATGACACCCAGCTGTGCTCTGGATGGTTTACCAAGTCCAAAT 250

QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 DB 251 TGCATGGTTTACTTTCAGGAAGCTGAGGAACCTGCTGCTGATGCCGAGCTCGAGTCTCAGTCT 310

QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 DB 311 TACGGAACAGGAGCCACCTGGGATCTATCTCAGTTTAAAGGAAGCCAGCACCATAGCA 370

QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 DB 371 GAGTACATAGTGGCTATCAGAGAGCCAGCCGATATGATCGGCTGCGACGACCCACAG 430

QY 101 LysArgGlnTrpGlnTrpIleAspGlyAlaMetTyrLysTrpArgSerTrpSerGly 120
 DB 431 AAGAGGAGCAGTGGGAGTGGATCGATCGATGGGCCATGTATCTGTACAGATCTCTGCTGGC 490

QY 121 LysSerMetGly-GlyAsnLysHisCysAlaGluMetSerSerAsnAsnPhenLeuT 140
 DB 491 AAGTTCATGGGCTGGGAACAAGCACTGTGCTGAGATGAGCTCCAAATAACAACCTTTACTAA 550

QY 140 hTrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
 DB 551 CTTGGAGAGCAGCAAGATGCAACAGCCCACTTCTGTGCAAGTACCCGACCA 606

RESULT 9
 AW957953 664 bp mRNA linear EST 01-JUN-2000
 LOCUS EST370023 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW957953
 VERSION AW957953.1 GI:18147636
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 664)
 REFERENCE Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt

I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J.
 Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
 Unpublished
 Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@igr.org
 Plate: 113
 Seq primer: Reverse.
 Location/Qualifiers
 1. .664
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGE resequences, MAGE"
 /note="Vector: pBluescriptSkm"

BASE COUNT	175 a	160 c	173 g	156 t
ORIGIN				

 Alignment Scores:
 Pred. No.: 6.08e-85 Length: 664
 Score: 834.00 Matches: 152
 Percent Similarity: 97.47% Conservative: 2
 Best Local Similarity: 96.20% Mismatches: 4
 Query Match: 94.99% Indels: 0
 DB: 10 Gaps: 0

US-09-525-041-2 (1-158) x AW957953 (1-664)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 DB 128 ATGGCTTCCAGAGAGCTGCGGCTGCTCTATATGCTGAGCTGCGCCAAACAGGAGTC 187

QY 21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 DB 188 CTGGGTGATATCATCATGAGACCCAGCTGCTCTCTGGATGGTTTACCAAGTCCAAAT 247

QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 DB 248 TGCATGGTTTACTTTCAGGAAGCTGAGGAACCTGCTGATGCCGAGCTCGAGTCTCAGTCT 307

QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 DB 308 TACGGAACAGGAGCCACCTGGGATCTATCTCAGTTTAAAGGAAGCCAGCACCATAGCA 367

QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 DB 368 GAGTACATAGTGGCTATCAGAGAGCCAGCCGATATGATGGCTGCGACCCACAG 427

QY 101 LysArgGlnTrpGlnTrpIleAspGlyAlaMetTyrLysTrpArgSerTrpSerGly 120
 DB 428 AAGAGGAGCAGTGGGAGTGGATCGATGGGCCATGTATCTGTACAGATCTCTGCTGGC 487

QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnPhenLeuThr 140
 DB 488 AAGTCCATGGGTTGGGAACAAGCACTGTGCTGAGATGAGCTCCCAATAACAACCTTTTAACT 547

QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
 DB 548 TGGAGACCAACGCACTTGCACCAAGCCCAACTTCTCTGTGAAGTACCCACCA 601

RESULT 10
 BG253845
 LOCUS BG253845
 DEFINITION 602366542F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:447491 5', mRNA sequence.
 ACCESSION BG253845
 VERSION BG253845.1 GI:12763661
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 908)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LAM10299 row: c column: 12
 High quality sequence stop: 734.
 FEATURES
 source
 1. 908
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4474691"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_91"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 254 a 219 c 234 g 201 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.29e-83 Length: 908
 Score: 822.00 Matches: 152
 Percent Similarity: 96.20% Conservatives: 0
 Best Local Similarity: 96.20% Mismatches: 6
 Query Match: 93.62% Indels: 1
 DB: 10 Gaps: 0
 US-09-525-041-2 (1-158) x BG253845 (1-908)
 Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 Db 136 ATGGCTTCAGAAAGCATGCGGCTGCTCTATTGCTGAGCTGCTGGCCAAACAGGAGTC 195
 Qy 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTTPheTyrHisLysSerAsn 40
 Db 196 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTCTGATGGTTTACCACAAAGTCCAAT 255
 Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 Db 256 TGCTATGGTTACTTCAGAAAGCTAGGAAGCTGGTCTGATGCGGAGCTCGAGTGTCTCT 315
 Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 Db 316 TACGGAAACGGAGGCCACCTGGCATCTATCTGATGTTTAAAGGAAGCCAGCACCATGACA 375
 Qy 81 GluTyrIleSerGlyTyrGlnArgSerClnProIleTyrIleGlyLeuHisAspProGln 100
 Db 376 GAGTACATAAGTGGCTATCAGAAAGCCAGCCGATGATGGCTGCGACGCCACAG 435
 Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 Db 436 AAGAGGCGAGCAGTGGCAGTGGATGATGGGGCCATGTATCTGTACAGATCTCTGTCTGGC 495
 Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnPheLeuThr 140
 Db 496 AAGTCCATGGGTGGGAACAAGCACTGTGCTGAGATGAGCTCCAATAACAACATTTTAAAC 555

Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
 Db 556 TTGGAGCAGCAAGATGCAACAGGCGCAACTCTCTGTGCAG-TACCGACCA 608
 RESULT 11
 BM788204 580 bp mRNA linear EST 05-MAR-2002
 LOCUS K-EST0067322 S3SNU16 Homo sapiens cDNA clone S3SNU16-34-B05 5',
 DEFINITION mRNA sequence.
 ACCESSION BM788204
 VERSION BM788204.1 GI:19136436
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 580)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 34 row: B column: 05
 High quality sequence stop: 580.
 FEATURES
 Location/Qualifiers
 1..580
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S3SNU16-34-B05"
 /sex="F"
 /tissue_type="Ascites"
 /cell_type="Lymphoblast-like"
 /cell_line="SNU-16"
 /lab_host="Top10F"
 /clone_lib="S3SNU16"
 /note="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;
 Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dT-selected mRNA by
 priming with dT-tailed vector. The dT-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."
 BASE COUNT 156 a 132 c 165 g 127 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.15e-80 Length: 580
 Score: 792.00 Matches: 144
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 90.21% Indels: 0
 DB: 12 Gaps: 0
 US-09-525-041-2 (1-158) x BM788204 (1-580)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 Db 147 ATGGCTTCCAGAGCATGGCGCTGCTCTATTGCTAGCTGCTGGCCAAACAGGAGTC 206
 Qy 21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 Db 207 CTGGGTGATATCATCATGAGACCCAGCTGTCTCTGGATGGTGTATACCAAGTCCAAT 266
 Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 Db 267 TGCTATGTTTACTTTCAGGAGCTGAGGAACCTGGTCTGATGCCGAGCTCGAGTGTCACTCT 326
 Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 Db 327 TACGGAACCGAGGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 386
 Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 Db 387 GAGTACATAGTGGCTATCAGAGAACCCAGCCGATATGGATGGCTTGCACGCCACAG 446
 Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 Db 447 AAGAGGCAGCAGTGGCAGTGGATTGATGGGCCATGATCTGTACAGATCTCGTGTCTGGC 506
 Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnLeuThr 140
 Db 507 AAGTCCATGGGTGGAAACAAGCACTGTGCTGAGATGAGTCCCAATACAACTTTTAACT 566
 Qy 141 TrpSerSerAsn 144
 Db 567 TGGAGCAGCAAC 578

RESULT 12
 AA314779
 LOCUS
 DEFINITION EST186601 Colon carcinoma (HCC) cell line II Homo sapiens cDNA 5'
 end, mRNA sequence.

ACCESSION
 VERSION AA314779
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS
 Adams M.D., Kerlavage A.R., Fleischmann R.D., Fuldner R.A., Bult
 C.J., Lee N.H., Kirkness E.F., Weinstock K.G., Gocayne J.D., White
 O., Sutton G., Blake J.A., Brandon R.C., Man-Wai C., Clayton R.A.,
 Cline T.R., Cotton M.D., Earle-Hughes J., Fine L.D., Fitzgerald
 L.M., Fitzhugh W.M., Fritchman J.L., Geoghagen N.S., Glodek A.,
 Gnehm C.L., Hanna M.C., Hedblom E., Hinkle P.S.Jr., Kelley J.M.,
 Kelley J.C., Liu L.-I., Marmaros S.M., Merrick J.M.,
 Moreno-Palauques R.F., McDonald L.A., Nguyen D.T., Pelligrino S.M.,
 Phillips C.A., Ryder S.E., Scott J.B., Saudek D.M., Shirley R.,
 Small K.V., Sprigge T.A., Utterback T.R., Weidman J.F., Li Y.,
 Bednarek D.P., Cao L., Cepeda M.A., Coleman T.A., Collins E.J.,
 Dimke D., Feng D.-F., Ferris A., Fischer C., Hastings G.A., He W.W.,
 Hu J.S., Greene J.M., Gruber J., Hudson P., Kim A.K., Kozak D.L.,
 Kunsch C., Hungjun J., Li H., Meissner P.S., Olsen H., Raymond L.,
 Wei Y.F., Wing J., Xu C., Yu G.L., Ruben S.M., Dillion P.J., Fannon
 M.R., Rosen C.A., Haseltine W.A., Fields C., Fraser C.M. and
 Venter J.C.

Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)
 96026280
 7566098
 Other ESTs: THC177381
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

FEATURES
 source
 1. .551
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="ATCC (inhost):111339"
 /db_xref="taxon:9606"
 /tissue_type="colon"
 /cell_line="KM12C"
 /cell_line="KM12C(HCC)-parental human colon carcinoma
 ;Dukes B2"
 /clone_lib="Colon carcinoma (HCC) cell line II"
 /note="Organ: colon; Vector: pBluescript SK-; Site_1:
 EcorI; Site_2: XhoI"
 BASE COUNT 148 a 125 c 153 g 124 t 1 others

Alignment Scores:
 Pred. No.: 1,846-79 Length: 551
 Score: 785.00 Matches: 143
 Percent Similarity: 99.31% Conservative: 0
 Best Local Similarity: 99.31% Mismatches: 1
 Query Match: 89.41% Indels: 0
 DB: 9 Gaps: 0

US-09-525-041-2 (1-158) x AA314779 (1-551)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 Db 118 ATGGCTTCCAGAGCATGGCGCTGCTCTATTGCTAGCTGCTGGCCAAACAGGAGTC 177
 Qy 21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 Db 178 CTGGGTGATATCATCATGAGACCCAGCTGTCTCTGGATGGTGTATACCAAGTCCAAT 237
 Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 Db 238 TGCTATGTTTACTTTCAGGAGCTGAGGAACCTGGTCTGATGCCGAGCTCGAGTGTCACTCT 297
 Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 Db 298 TACGGAACCGAGGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 357
 Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 Db 358 GAGTACATAAGTGGCTATCAGAGAACCCAGCCGATATGGATGGCTGGCCAGCCACAG 417
 Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 Db 418 AAGAGGCAGCAGTGGCAGTGGATTGATGGGCCATGATCTGTACAGATCTCGTGTCTGNC 477
 Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnLeuThr 140
 Db 478 AAGTCCATGGGTGGGAACAAGCACTGTCTGAGATGAGTCCCAATACAACTTTTAACT 537
 Qy 141 TrpSerSerAsn 144
 Db 538 TGGAGCAGCAAC 549

RESULT 13
 BG256113

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

BG256113 602367238F1 NTH_MGC_91 Homo sapiens cDNA clone IMAGE:4475648 5',
 mRNA sequence.
 BG256113
 BG256113.1 GI:12765929
 EST.
 Homo sapiens (human)

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 968)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM0301 row: k column: 09
 High quality sequence stop: 672.
 Location/Qualifiers
 source
 1..968
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4475648"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 91"
 /notes="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC Library."
 268 a 235 c 245 g 220 t

BASE COUNT 268 a 235 c 245 g 220 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,28e-79 Length: 968
 Score: 785.00 Matches: 141
 Percent Similarity: 99.30% Conservative: 0
 Best Local Similarity: 99.30% Mismatches: 1
 Query Match: 89.41% Indels: 1
 DB: 10 Gaps: 0

US-09-525-041-2 (1-158) x BG256113 (1-968)

Qy 17 LysThrGlyValLeuGlyAspIleIleMetArgProSerCysAlaProGlyTyrPheTyr 36
 Db 1 AAAACAGGAGT-CTGGGAGATCATCATGAGACCCAGCTGTCTCTGGATGTTTAC 59
 Qy 37 HisLysSerAsnCysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAlaGluLeu 56
 Db 60 CACAGTCCANTTCTATGTTTCTTACGAGGCTGAGGAAGCTGGTCTGATGCCGAGCTC 119
 Qy 57 GluCysGlnSerTyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuGlyGluAla 76
 Db 120 GAGTGTCACTTACCGAAACGGAGCCACCTGGCATCTATCTCTGAGTTTAAAGGAAGCC 179
 Qy 77 SerThrIleAlaGluTyrIleSerGlyTyrGlnArgSerGlnProIleTyrIleGlyLeu 96
 Db 180 AGCACCATAGCAGAGTACATAGTGGCTATCAGAGAAGCCAGCCGATATGATGGCTG 239
 Qy 97 HisAspProGlnLysArgGlnTyrGlnTrpIleAspGlyAlaMetTyrLeuTyrArg 116
 Db 240 CACGACCCACAGAGAGCAGCAGTGGCAGTGGATGATGGGGCCATGATCTGTACAGA 299
 Qy 117 SerTrpSerGlyLysSerMetGlyCysAsnLysHisCysAlaGluMetSerSerAsnAsn 136
 Db 300 TCTTGGTCTGGCAAGTCCATGGGTGGGAACAAGCACTGTGTGAGATGAGCTCCAAATAC 359
 Qy 137 AsnPheLeuThrTrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyr 156
 Db 360 AACTTTTTTAATTGAGCAGCAACGAATGCAACAAGGCCAACACTTCTCTGTGCAAGTAC 419

Qy 157 ArgPro 158
 Db 420 CGACCA 425
 RESULT 14
 AA581222
 LOCUS
 DEFINITION
 AA581222 440 bp mRNA linear EST 05-JAN-1998
 similar to SW:LECG_CROAT P21963 GALACTOSE-SPECIFIC LECTIN. ;, mRNA
 sequence.
 ACCESSION
 VERSION AA581222.1 GI:2358994
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 440)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies Inc., David Krizman,
 Ph.D.
 CDNA Library Arraying: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -28m13 rev1 ET from Amersham
 High quality sequence stop: 410.
 Location/Qualifiers
 source
 1..440
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:802600"
 /tissue_type="bulk tumor"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Col"
 /notes="Organ: colon; Vector: pCMV-SPORT2; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr."
 125 a 105 c 114 g 96 t
 BASE COUNT 125 a 105 c 114 g 96 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,31e-77 Length: 440
 Score: 761.00 Matches: 133
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.67% Indels: 0
 DB: 9 Gaps: 0
 US-09-525-041-2 (1-158) x AA581222 (1-440)
 Qy 26 MetArgProSerCysAlaProGlyTyrPheTyrHisLysSerAsnCysTyrGlyTyrPhe 45
 Db 10 ATGAGACCCAGCTGTCTCTCTGGATGTTTACCAAGTCCCAATGCTATGTTACTTC 69
 Qy 46 ArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSerTyrGlyAsnGlyAla 65
 Db 70 AGGAACCTGAGGAACCTGGTCTGATGCCGAGCTCGAGTGTGAGTTCACGAAACGGAGCC 139
 Qy 66 HisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAlaGluTyrIleSerGly 85
 Db 130 CACCTGGCATCTATCTCTGAGTTTAAAGGAAGCCAGCACCATACAGTACATAGTGC 189
 Qy 86 TyrGlnArgSerGlnProIleTyrIleGlyLeuHisAspProGlnLysArgGlnTyr 105

Db	190	TATCAGAGAACCCAGCCGATATGGAATTGGCTTGCAGCACCACACAGAGAGGCGAGCAGTGG	249
Qy	106	GlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGlyLysSerMetGlyGly	125
Db	250	CAGTGGATTGATGGGCCATGTATCTGTACAGATCCTGGTCTGGCAAGTCCATGGGTGGG	309
Qy	126	AsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThrTrpSerSerAsnGlu	145
Db	310	AACAAGCACCTGTGCTGAGATGAGCTCCAATAACAACATTTTAACTTGGAGCAGCAACGAA	369
Qy	146	CysAsnLysArgGlnHisPheLeuCysLysTyrArgPro	158
Db	370	TGCAACAAGGGCCAAACATCTCTGTGCAAGTATCCGACCA	408

RESULT 15	
AI791498	526 bp mRNA linear EST 13-DEC-1999
LOCUS	
DEFINITION	ni23g10.y5 NCI_CGAP Co4 Homo sapiens cDNA clone IMAGE:968898 5', similar to SW:DECG_CROAT P21963 GALACTOSE-SPECIFIC LECTIN. ;, mRNA sequence.
ACCESSION	AI791498
VERSION	AI791498.1 GI:5339214
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1. (bases 1 to 526)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished
COMMENT	Other ESTs: ni23g10.s1 Contact: Robert Strausberg, Ph.D. Email: cgapbs-x@mail.nih.gov Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. cDNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

This read is a RESEQUENCE of a previously sequenced human clone
Original clone citation: see original entry for original citation
information
This 5' resequenced clone has no previous 5' data to verify this
new read against
Insert Length: 1216 Std Error: 0.00
Seq primer: ~40RP from Gibco
High quality sequence stop: 460.

BASE COUNT	141 a	121 c	149 g	115 t
ORIGIN				

Alignment Scores:

Pred. No.:	1,028-75	Length:	526
Alignment Scores:		Matches:	137
Score:	752.00	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	85.65%	Gaps:	0
DB:	9		

US-09-525-041-2 (1-158) x AI791498 (1-526)

[illegible]

Search completed: December 31, 2003, 10:39:40
Job time : 2062 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 31, 2003, 09:11:17 ; Search time 71 Seconds
(without alignments)
353.223 Million cell updates/sec

Title: US-09-525-041-2
Perfect score: 878
Sequence: 1 MASRMRLLLLLSCLAKTV.....LTWSSNECKRQHFLCKYRP 158

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	878	100.0	158	AAW12691	Human colon specif
2	878	100.0	158	AAW37929	A human Reg I-gamm
3	878	100.0	158	AAW37866	Human protein comp
4	878	100.0	158	AAW84274	Protein encoded by
5	878	100.0	158	AAW12900	Human colon specif
6	878	100.0	158	AAW24517	CSI-152 clone pred
7	878	100.0	158	AAW24518	C880P similar amin
8	878	100.0	158	AAW24519	C880P similar amin
9	878	100.0	158	AAW24520	C880P similar amin

10	878	100.0	158	22	AAW24521	C880P similar amin
11	878	100.0	158	22	AAW74934	Human TSA7005 prot
12	878	100.0	158	23	AAE29829	Human REG-like pro
13	878	100.0	158	23	ABB78993	Human Reg IV prote
14	878	100.0	158	24	ABP56022	Human REG-like pro
15	878	100.0	158	24	ABP55366	Human colon tumour
16	878	100.0	158	24	ABP55367	Human colon specif
17	878	100.0	158	24	ABP55368	Human colon specif
18	878	100.0	158	24	ABP55369	Human colon specif
19	878	100.0	158	24	ABP55370	Human colon specif
20	878	100.0	166	22	AAW75620	Human colon cancer
21	867	98.7	158	23	ABP99310	Amino acid sequenc
22	856	97.5	153	21	AAI92267	Human cancer speci
23	695	79.2	122	24	ABP76307	Human GENSET prote
24	307.5	35.0	134	23	ABP59097	Human zinc finger
25	254.5	29.0	165	9	AAW81513	Sequence encoded b
26	254.5	29.0	165	9	AAW81388	Sequence encoded b
27	254.5	29.0	165	14	AAW34535	MUREG-1. Mus musc
28	254.5	29.0	165	15	AAW59289	Rat reg protein.
29	252.5	28.8	165	10	AAW94615	Rat reg protein.
30	245	27.9	144	15	AAW66595	Rat reg protein (G
31	245	27.9	146	15	AAW66594	Rat reg protein (G
32	243	27.7	133	15	AAW66596	Rat reg protein (I
33	240	27.3	158	17	AAW18565	Coral cobra venom
34	240	27.3	174	21	AAW34301	Human pancreatic c
35	236.5	26.9	173	14	AAW34536	MUREG-2. Mus musc
36	229	26.1	175	11	AAW05904	Bovine pancreatic
37	225.5	25.7	166	15	AAW59288	Human reg protein.
38	225.5	25.7	166	22	AAW71653	Human colon associ
39	225.5	25.7	166	22	AAW71666	Human colon associ
40	225.5	25.7	174	21	AAW37377	Human cancer assoc
41	224	25.5	144	15	AAW66592	Human reg protein
42	224	25.5	146	15	AAW66591	Human reg protein
43	224	25.5	147	11	AAW06425	Reg protein analog
44	223	25.4	175	23	ABJ10605	Human novel protei
45	222	25.3	133	15	AAW66593	Human reg protein

ALIGNMENTS

RESULT 1
AAW12691
ID AAW12691 standard; Protein; 158 AA.
XX
AC AAW12691;
XX
DT 31-MAY-1997 (first entry)
XX
DE Human colon specific protein.
XX
KW Colon specific protein; colon cancer; metastasis; diagnosis;
XX therapy; antibody; vaccine; agonist; antagonist.
XX
OS Homo sapiens.
XX
PN WO9639541-A1.
XX
PD 12-DEC-1996.
XX
PF 06-JUN-1995; 95WO-US07169.
XX
PR 06-JUN-1995; 95WO-US07169.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Dillon PJ, Li Y, Soppet DR;
XX
DR WPI; 1997-043162/04.
XX
PT N-PSDB; AAT51784.
XX
PT New isolated colon specific gene - used to develop prods. for use in
the diagnosis and treatment of colon disorders, partic. colon

PT cancer.
 XX Claim 1; Page 53; 64pp; English.
 XX
 CC A human colon specific protein (AAW12691) is a potential diagnostic
 CC marker for colon cancer. It is believed that the presence of
 CC active transcription of the colon specific gene in non-colon cells
 CC of a host is indicative of the colon cancer metastases. The amino acid
 CC sequence of the colon specific protein was deduced from a cDNA
 CC clone (AAT51784) isolated from a human colon cancer cDNA library.
 CC Recombinant colon specific protein can be produced in transformed
 CC host (e.g. bacterial, insect) cells and used to develop prods. for
 CC the diagnosis and treatment of colon disorders, partic. colon
 CC cancer metastasis. Antibodies raised against the protein can be
 CC used to target cancer cells and as part of a colon cancer vaccine.
 XX
 SQ Sequence 158 AA;
 Query Match 100.0%; Score 878; DB 18; Length 158;
 Best Local Similarity 100.0%; Pred. No. 1.3e-84;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWYHKSNICYGFRKLRNWSDALEECQS 60
 DB 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWYHKSNICYGFRKLRNWSDALEECQS 60
 QY 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKROQWIDGAMLYRSWSG 120
 DB 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKROQWIDGAMLYRSWSG 120
 QY 121 KSMGNGKHCHEMSSNNFLTWSSNECNKROHFLCKYRP 158
 DB 121 KSMGNGKHCHEMSSNNFLTWSSNECNKROHFLCKYRP 158
 RESULT 2
 AAW37929
 ID AAW37929 standard; Protein; 158 AA.
 XX
 AC AAW37929;
 XX
 DT 21-AUG-1998 (first entry)
 XX
 DE A human Reg I-gamma protein.
 XX
 KW Human; Reg I-gamma protein; C-type lectin; Incyte clone 1310334;
 KW regulation; cell growth; development; tumorigenesis; neurodegeneration;
 KW inhibition; treatment; prevention; neoplasia; metastasis;
 KW neurodegenerative change; Alzheimer's disease; Down's syndrome;
 KW regeneration; pancreatic beta-cells; diabetes.
 XX
 OS Homo sapiens.
 XX
 FN WO9816640-A1.
 XX
 PD 23-APR-1998.
 XX
 PF 07-OCT-1997; 97WO-US18174.
 XX
 PR 11-OCT-1996; 96US-0729103.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Bandman O, Goli SK;
 XX
 DR WPI; 1998-251287/22.
 XX
 DR N-PSDB; AAV29156.
 XX
 PT New isolated human Reg I-gamma protein - useful for developing
 PT products for treating, e.g. diabetes, tumours or neuro-degenerative
 PT disease such as Alzheimer's
 XX
 PS Claim 1; Fig 1; 72pp; English.

XX
 CC The present sequence represents a human Reg I-gamma protein, which
 CC comprises a C-type lectin. The sequence was identified in incyte clone
 CC 1310334. Reg I-gamma protein is involved in regulation of cell growth
 CC and development. Since the overexpression of reg proteins is associated
 CC with tumorigenesis and neurodegeneration, inhibition of human Reg
 CC I-gamma expression can be used for treating or preventing neoplasia or
 CC metastasis and neurodegenerative changes associated with Alzheimer's
 CC disease and other disorders of the central nervous system, e.g. Down's
 CC syndrome. Reg I-gamma can also be used in therapeutics to induce
 CC regeneration of pancreatic beta-cells in the treatment of diabetes. The
 CC products can also be used for detection for, e.g. expression of REG
 CC I-gamma, diagnosis and drug screen
 XX
 SQ Sequence 158 AA;
 Query Match 100.0%; Score 878; DB 19; Length 158;
 Best Local Similarity 100.0%; Pred. No. 1.3e-84;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWYHKSNICYGFRKLRNWSDALEECQS 60
 DB 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWYHKSNICYGFRKLRNWSDALEECQS 60
 QY 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKROQWIDGAMLYRSWSG 120
 DB 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKROQWIDGAMLYRSWSG 120
 QY 121 KSMGNGKHCHEMSSNNFLTWSSNECNKROHFLCKYRP 158
 DB 121 KSMGNGKHCHEMSSNNFLTWSSNECNKROHFLCKYRP 158
 RESULT 3
 AAW37866
 ID AAW37866 standard; Protein; 158 AA.
 XX
 AC AAW37866;
 XX
 DT 10-AUG-1998 (first entry)
 XX
 DE Human protein comprising secretory signal amino acid sequence 3.
 XX
 KW Human protein; secretory signal; nutritional source; cytokine;
 KW immunity; haematopoiesis; activin; inhibin; tumour; chemotactic;
 KW chemokinetic; thrombolytic; anti-inflammatory; inhibition;
 KW stomach cancer cell.
 XX
 OS Homo sapiens.
 XX
 FN WO9811217-A2.
 XX
 PD 19-MAR-1998.
 XX
 PF 12-SEP-1997; 97WO-JP03239.
 XX
 PR 13-SEP-1996; 96JP-0243060.
 XX
 PA (PROT-) PROTEGENE INC.
 XX
 PA (SAGA) SAGAMI CHEM RES CENTRE.
 XX
 PI Kato S, Kobayashi M, Sekine S, Yamaguchi T;
 XX
 DR WPI; 1998-207380/18.
 XX
 DR N-PSDB; AAV29035, AAV29036.
 XX
 PT Human proteins with secretory signal sequences - used to treat
 PT immune deficiencies, infections, tumours, and haematopoietic
 PT disorders, etc.
 XX
 PS Claim 1; Pages 67-68; 131pp; English.

CC This is the amino acid sequence of a novel human protein comprising
CC a secretory signal isolated from stomach cancer cells. Its proteins
CC can be used as nutritional sources or supplements. The proteins may
CC also have cytokine functions, immune modulating functions,
CC haematopoiesis regulating activity, activin/inhibin regulating
CC activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity.

XX Sequence 158 AA;
SQ

Query Match 100.0%; Score 878; DB 19; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.3e-84;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRMRLLLLLCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFKLRNWSDALECOS 60
DB 1 MASRMRLLLLLCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFKLRNWSDALECOS 60

QY 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQWQWIDGAMLYRSWSG 120
DB 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQWQWIDGAMLYRSWSG 120

QY 121 KSMGNGKHCAEMSSNNFLTWSSNECNKQHFCKYRP 158
DB 121 KSMGNGKHCAEMSSNNFLTWSSNECNKQHFCKYRP 158

RESULT 4
AAW84274
ID AAW84274 standard; Protein; 158 AA.
XX
AC AAW84274;
XX
DT 25-MAR-1999 (first entry)
XX
DE Protein encoded by a human colon specific gene.
XX
KW Human; colon specific gene; diagnosis; colon disorder; colon cancer;
KW viability; colon cancer cell.
XX
OS Homo sapiens.
XX
PN US5861494-A.
XX
PD 19-JAN-1999.
XX
PF 06-JUN-1995; 95US-0468413.
XX
PR 06-JUN-1995; 95US-0468413.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Dillon PJ, Li Y, Soppet DR;
XX
DR WPI; 1999-130432/11.
DR N-PSDB; AAX03195.
XX
PT Isolated human colon specific gene - used to develop products for
PT the diagnosis and treatment of disorders of the colon, e.g. colon
PT cancer and metastases
XX
PS Claim 1; Fig 1A-C; 20pp; English.
XX
CC The present sequence is encoded by a human colon specific gene. The
CC nucleic acid sequence can be used to develop products for the diagnosis
CC of a disorder of the colon, e.g. colon cancer or metastases. The
CC products can also be used to screen for agonists or antagonists for
CC the polypeptides. The antagonists may be used to treat colon cancer,
CC since they interact with the function of colon specific polypeptides
CC to inhibit functions which are necessary for the viability of colon
CC cancer cells. The products can also be used for the production of
CC antibodies and for the identification of receptors for the

CC polypeptides.
XX Sequence 158 AA;
SQ

Query Match 100.0%; Score 878; DB 20; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.3e-84;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRMRLLLLLCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFKLRNWSDALECOS 60
DB 1 MASRMRLLLLLCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFKLRNWSDALECOS 60

QY 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQWQWIDGAMLYRSWSG 120
DB 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQWQWIDGAMLYRSWSG 120

QY 121 KSMGNGKHCAEMSSNNFLTWSSNECNKQHFCKYRP 158
DB 121 KSMGNGKHCAEMSSNNFLTWSSNECNKQHFCKYRP 158

RESULT 5
AAB12900
ID AAB12900 standard; Protein; 158 AA.
XX
AC AAB12900;
XX
DT 14-NOV-2000 (first entry)
XX
DE Human colon specific protein sequence.
XX
KW Human; colon specific; colon cancer; metastasis; diagnose; treatment;
KW cytostatic.
XX
OS Homo sapiens.
XX
PN US6080722-A.
XX
PD 27-JUN-2000.
XX
PF 29-SEP-1998; 98US-0162508.
XX
PR 06-JUN-1995; 95US-0468413.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Dillon PJ, Soppet DR, Li Y;
XX
DR WPI; 2000-464055/40.
DR N-PSDB; AAA62951.
XX
PT Novel human colon specific polypeptides and polynucleotides for
PT diagnosis and treatment of colon cancer, for screening compounds which
PT interact with polypeptide, for synthesis of DNA and manufacture of DNA
PT vectors
XX
PS Claim 1; Fig 1; 20pp; English.
XX
CC This invention relates to a purified human protein, which is primarily
CC expressed in tissue derived from the colon. The protein is 152 amino
CC acids in length and exhibits cytostatic activity. The present sequence
CC represents the amino acid sequence of the colon specific protein. The
CC protein can be used in the diagnosis and treatment of colon cancer, and
CC it is thought that abnormally high levels of the gene expression in
CC non-colon cells is an indication of colon cancer metastasis.

XX Sequence 158 AA;
SQ

Query Match 100.0%; Score 878; DB 21; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.3e-84;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRMRLLLLLCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFKLRNWSDALECOS 60

Db 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALECSQ 60
Qy 61 YGNGAHLASILSLKEASTIAEYISGYQORSQPIWIGLHDPKQKQOWIDGAMLYRWSWG 120
Db 61 YGNGAHLASILSLKEASTIAEYISGYQORSQPIWIGLHDPKQKQOWIDGAMLYRWSWG 120
Qy 121 KSMGKNKHCAMSSNNNFLTWSNNECNKQHFCLKYRP 158
Db 121. KSMGKNKHCAMSSNNNFLTWSNNECNKQHFCLKYRP 158
RESULT 6
ID AAM24517
XX AAM24517 standard; Protein; 158 AA.
AC AAM24517;
XX
DT 12-OCT-2001 (first entry)
XX
DE CSI-152 clone predicted amino acid sequence.
XX
KW Human; immunotherapy; diagnosis; colon cancer; colon tumour;
KW immunogenic; gene therapy; vaccine; colonic cancer.
XX
OS Homo sapiens.
XX
PN WO200149716-A2.
XX
PD 12-JUL-2001.
XX
PF 29-DEC-2000; 2000WO-US35596.
XX
PR 30-DEC-1999; 99US-0476296.
PR 10-JAN-2000; 2000US-0480321.
PR 15-FEB-2000; 2000US-0504629.
PR 06-MAR-2000; 2000US-0519444.
PR 19-MAY-2000; 2000US-0575251.
PR 29-JUN-2000; 2000US-0609448.
PR 28-AUG-2000; 2000US-0649811.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;
XX
DR WPI; 2001-441847/47.
XX
CC Colon tumor associated proteins and nucleic acids useful for the
CC prevention, diagnosis and treatment of colonic cancer -
PS Claim 2; Page 463; 472pp; English.
XX
CC The present invention describes colon tumour associated proteins (I) and
CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be
CC used to treat disorders associated with decreased expression by
CC rectifying mutations or deletions in a patient's genome that affect the
CC activity of TCAPs by expressing inactive proteins or to supplement the
CC patients own production of them. Additionally, (II) may be used to
CC produce the TCAP proteins, by inserting the nucleic acids into a host
CC cell culturing the cell to express the protein. (II) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC polymerase chain reaction (PCR) and hybridisation assays to detect and
CC quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. (I) may
CC also be used as antigens in the production of antibodies against TCAPs
CC and in assays to identify modulators of TCAP expression and activity.
CC Anti-(I) antibodies and antagonists may also be used to down regulate
CC TCAP expression and activity. The anti-(I) antibodies may also be used

CC as diagnostic agents for detecting the presence of TCAPs in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)). AAI28460 to AAI29512
CC and AAM24494 to AAM24523 represent nucleotide and amino acid sequences
CC given in the exemplification of the present invention.
XX
SQ Sequence 158 AA;
Query Match 100.0%; Score 878; DB 22; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.3e-84;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALECSQ 60
Db 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDALECSQ 60
Qy 61 YGNGAHLASILSLKEASTIAEYISGYQORSQPIWIGLHDPKQKQOWIDGAMLYRWSWG 120
Db 61 YGNGAHLASILSLKEASTIAEYISGYQORSQPIWIGLHDPKQKQOWIDGAMLYRWSWG 120
Qy 121 KSMGKNKHCAMSSNNNFLTWSNNECNKQHFCLKYRP 158
Db 121 KSMGKNKHCAMSSNNNFLTWSNNECNKQHFCLKYRP 158
RESULT 7
ID AAM24518
XX AAM24518 standard; Protein; 158 AA.
AC AAM24518;
XX
DT 12-OCT-2001 (first entry)
XX
DE C880P similar amino acid sequence (GENESEQ W12691).
XX
KW Human; immunotherapy; diagnosis; colon cancer; colon tumour;
KW immunogenic; gene therapy; vaccine; colonic cancer.
XX
OS Homo sapiens.
XX
PN WO200149716-A2.
XX
PD 12-JUL-2001.
XX
PF 29-DEC-2000; 2000WO-US35596.
XX
PR 30-DEC-1999; 99US-0476296.
PR 10-JAN-2000; 2000US-0480321.
PR 15-FEB-2000; 2000US-0504629.
PR 06-MAR-2000; 2000US-0519444.
PR 19-MAY-2000; 2000US-0575251.
PR 29-JUN-2000; 2000US-0609448.
PR 28-AUG-2000; 2000US-0649811.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;
XX
DR WPI; 2001-441847/47.
XX
CC Colon tumor associated proteins and nucleic acids useful for the
CC prevention, diagnosis and treatment of colonic cancer -
PS Claim 2; Page 467; 472pp; English.
XX
CC The present invention describes colon tumour associated proteins (I) and
CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be
CC used to treat disorders associated with decreased expression by
CC rectifying mutations or deletions in a patient's genome that affect the

CC activity of TCAPs by expressing inactive proteins or to supplement the
 CC patients own production of them. Additionally, (ii) may be used to
 CC produce the TCAP proteins, by inserting the nucleic acids into a host
 CC cell culturing the cell to express the protein. (ii) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC polymerase chain reaction (PCR) and hybridisation assays to detect and
 CC quantitate the presence of similar nucleic acids in samples, and
 CC therefore which patients may be in need of restorative therapy. (i) may
 CC also be used as antigens in the production of antibodies against TCAPs
 CC and in assays to identify modulators of TCAP expression and activity.
 CC Anti-(i) antibodies and antagonists may also be used to down regulate
 CC TCAP expression and activity. The anti-(i) antibodies may also be used
 CC as diagnostic agents for detecting the presence of TCAPs in samples
 CC (e.g. by enzyme linked immunosorbant assay (ELISA)). AA128460 to AA129512
 CC and AA24494 to AA24523 represent nucleotide and amino acid sequences
 CC given in the exemplification of the present invention.
 XX Sequence 158 AA;
 SQ

Query Match 100.0%; Score 878; DB 22; Length 158;
 Best Local Similarity 100.0%; Pred. No. 1.3e-84;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYPRKLRNWSDALECOS 60
 DB 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYPRKLRNWSDALECOS 60
 OY 61 YNGAHLASILSLKEASTIAEYISGYQRSQPTWIGLHDPKQKQOWIDGAMLYRSWSG 120
 DB 61 YNGAHLASILSLKEASTIAEYISGYQRSQPTWIGLHDPKQKQOWIDGAMLYRSWSG 120
 OY 121 KSMGGNKHCAEMSSNNNFLTWSNECNKRQHFCLKYRP 158
 DB 121 KSMGGNKHCAEMSSNNNFLTWSNECNKRQHFCLKYRP 158
 RESULT 8
 AA24519
 ID AA24519 standard; Protein; 158 AA.
 AC AA24519;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE C880P similar amino acid sequence (GENESEQ W37866).
 XX
 DE Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 XX
 KW Human; immunogenic; gene therapy; vaccine; colonic cancer.
 KW
 OS Homo sapiens.
 XX
 PN WO200149716-A2.
 XX
 PD 12-JUL-2001.
 XX
 PF 29-DEC-2000; 2000WO-US35596.
 XX
 PR 30-DEC-1999; 99US-0476296.
 PR 10-JAN-2000; 2000US-0480321.
 PR 15-FEB-2000; 2000US-0504629.
 PR 06-MAR-2000; 2000US-0519444.
 PR 19-MAY-2000; 2000US-0575251.
 PR 29-JUN-2000; 2000US-0609448.
 PR 28-AUG-2000; 2000US-0649811.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
 PI King GE, Wang T, Jiang Y;
 XX
 XX WPI; 2001-441847/47.
 DR
 XX Colon tumor associated proteins and nucleic acids useful for the
 PT

PT prevention, diagnosis and treatment of colonic cancer -
 XX Claim 2; Page 467-468; 472pp; English.
 XX
 CC The present invention describes colon tumour associated proteins (i) and
 CC the polynucleotides (ii) that encode them. (i) have cytostatic activity,
 CC (i) and (ii) can be used in gene therapy and vaccine production. (i) and
 CC (ii) may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate colon tumour associated protein (TCAP)
 CC expression, such as colonic cancer. For example, (i) and (ii) may be
 CC used to treat disorders associated with decreased expression by
 CC rectifying mutations or deletions in a patient's genome that affect the
 CC activity of TCAPs by expressing inactive proteins or to supplement the
 CC patients own production of them. Additionally, (ii) may be used to
 CC produce the TCAP proteins, by inserting the nucleic acids into a host
 CC cell culturing the cell to express the protein. (ii) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC polymerase chain reaction (PCR) and hybridisation assays to detect and
 CC quantitate the presence of similar nucleic acids in samples, and
 CC therefore which patients may be in need of restorative therapy. (i) may
 CC also be used as antigens in the production of antibodies against TCAPs
 CC and in assays to identify modulators of TCAP expression and activity.
 CC Anti-(i) antibodies and antagonists may also be used to down regulate
 CC TCAP expression and activity. The anti-(i) antibodies may also be used
 CC as diagnostic agents for detecting the presence of TCAPs in samples
 CC (e.g. by enzyme linked immunosorbant assay (ELISA)). AA128460 to AA129512
 CC and AA24494 to AA24523 represent nucleotide and amino acid sequences
 CC given in the exemplification of the present invention.
 XX Sequence 158 AA;
 SQ

Query Match 100.0%; Score 878; DB 22; Length 158;
 Best Local Similarity 100.0%; Pred. No. 1.3e-84;
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYPRKLRNWSDALECOS 60
 DB 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYPRKLRNWSDALECOS 60
 OY 61 YNGAHLASILSLKEASTIAEYISGYQRSQPTWIGLHDPKQKQOWIDGAMLYRSWSG 120
 DB 61 YNGAHLASILSLKEASTIAEYISGYQRSQPTWIGLHDPKQKQOWIDGAMLYRSWSG 120
 OY 121 KSMGGNKHCAEMSSNNNFLTWSNECNKRQHFCLKYRP 158
 DB 121 KSMGGNKHCAEMSSNNNFLTWSNECNKRQHFCLKYRP 158
 RESULT 9
 AA24520
 ID AA24520 standard; Protein; 158 AA.
 AC AA24520;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE C880P similar amino acid sequence (GENESEQ W37929).
 DE
 DE Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 KW Human; immunogenic; gene therapy; vaccine; colonic cancer.
 KW
 OS Homo sapiens.
 XX
 PN WO200149716-A2.
 XX
 PD 12-JUL-2001.
 XX
 PF 29-DEC-2000; 2000WO-US35596.
 XX
 PR 30-DEC-1999; 99US-0476296.
 PR 10-JAN-2000; 2000US-0480321.
 PR 15-FEB-2000; 2000US-0504629.
 PR 06-MAR-2000; 2000US-0519444.
 PR 19-MAY-2000; 2000US-0575251.
 PR 29-JUN-2000; 2000US-0609448.
 PR 28-AUG-2000; 2000US-0649811.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
 PI King GE, Wang T, Jiang Y;
 XX
 XX WPI; 2001-441847/47.
 DR
 XX Colon tumor associated proteins and nucleic acids useful for the
 PT

XX 02-OCT-2001; 2001WO-US30732.
XX 02-OCT-2000; 2000US-237271P.
XX (FARB) BAYER CORP.
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
XX Thiaglingam A, Lewis ME;
XX WPI; 2002-426115/45.
XX N-PSDB; ABQ60776.
XX New isolated nucleic acid that is differentially expressed in cancer
XX tissues useful for determining the presence of colon cancer in a cell
XX or tissue type, and in antisense therapy -
XX Claim 5; Fig 3; 796pp; English.
XX
XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
XX expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
XX encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
XX used in antisense therapy. An antibody immunoreactive with a polypeptide
XX encoded by (I) is useful for detecting cancer in a patient sample, and
XX for detecting the presence or absence of a polynucleotide encoded by a
XX nucleic acid which hybridises to (I) in a cell. A probe/primer derived
XX from (I) can be used for determining the presence of a nucleic acid which
XX hybridises to (I), and for determining the phenotype of cells in a sample
XX of cells from a patient. (I) is useful for determining the presence of
XX colon cancer in a cell or tissue type, for determining the presence or
XX state of other type of cancer, in antisense therapy, to generate
XX macroarrays on a solid surface, to identify a chromosome on which the
XX corresponding gene resides, and in tissue profiling, forensics, genetic
XX analysis, mapping and diagnostic applications. (I) can be used to raise
XX antibodies, and to screen for peptide analogues and antagonists.
XX
XX Sequence 158 AA;
XX
Query Match 100.0%; Score 878; DB 23; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.3e-84;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWYHKSNICYGYFRKLRNWSDALEECQS 60
DB 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWYHKSNICYGYFRKLRNWSDALEECQS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
DB 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
QY 121 KSMGKNKHCAMSSNNNFLTWSNECNKQHFCLKYRP 158
DB 121 KSMGKNKHCAMSSNNNFLTWSNECNKQHFCLKYRP 158
RESULT 14
ABP56022
ID ABP56022 standard; Protein; 158 AA.
XX
XX ABP56022;
XX
XX 26-FEB-2003 (first entry)
XX
XX Human REG-like protein (REL P) SEQ ID NO:2.
XX
XX Human; REG-like protein; RELP; immunoglobulin derived protein; Ig;
XX immunoglobulin; cytostatic; Ig agonist; immunoglobulin agonist; cancer;
XX protein therapy; RELP human Ig derived protein; chromosome 1p12-13.1.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..26

FT Protein /label= signal
FT 27..158
FT /label= RELP
XX
XX WO200274916-A2.
XX
XX 26-SEP-2002.
XX
XX 14-MAR-2002; 2002WO-US07945.
XX
XX 16-MAR-2001; 2001US-276305P.
XX
XX (CENZ) CENTOCOR INC.
XX
XX Heiskala M;
XX
XX WPI; 2003-103204/09.
XX N-PSDB; ABZ21635.
XX
XX New isolated REG-like protein (REL P) human immunoglobulin derived
XX protein or specified portion or variant, useful for preventing or
XX treating a RELP protein mediated condition or malignant condition, e.g.
XX cancer -
XX Claim 1; Fig 2; 101pp; English.
XX
XX The present sequence represents a new isolated REG-like protein (REL P)
XX human immunoglobulin (Ig) derived protein. RELP comprises: (a) a human
XX variable and constant region; or (b) an isolated human Ig derived
XX protein or specified portion or variant encoded by a nucleic acid.
XX RELP has cytostatic activity and can be used as an Ig agonist and in
XX protein therapy. The RELP human Ig derived protein or a specified
XX portion or variant can be used for preventing or treating a RELP protein
XX mediated condition, malignant condition or disease condition, e.g.
XX cancer. The nucleic acids can be used in producing RELP Ig derived
XX protein. The human RELP protein of the present invention is located to
XX chromosome 1p12-13.1.
XX
XX Sequence 158 AA;
XX
Query Match 100.0%; Score 878; DB 24; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.3e-84;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWYHKSNICYGYFRKLRNWSDALEECQS 60
DB 1 MASRSMLLLLSCLAKTGVLDIIMRPSCAPGWYHKSNICYGYFRKLRNWSDALEECQS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
DB 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
QY 121 KSMGKNKHCAMSSNNNFLTWSNECNKQHFCLKYRP 158
DB 121 KSMGKNKHCAMSSNNNFLTWSNECNKQHFCLKYRP 158
RESULT 15
ABP55366
ID ABP55366 standard; Protein; 158 AA.
XX
XX ABP55366;
XX
XX 30-JAN-2003 (first entry)
XX
XX Human colon tumour protein for clone CS1-152 SEQ ID NO:1070.
XX
XX Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;
XX tumour; immune response; immunostimulant; cytostatic; vaccine.
XX
XX Homo sapiens.
XX
XX WO200283070-A2.

XX 24-OCT-2002.
PD
XX
XX 09-APR-2002; 2002WO-US11475.
PF
XX
XX 10-APR-2001; 2001US-0833263.
PR
XX 03-AUG-2001; 2001US-0922217.
PR
XX 19-DEC-2001; 2001US-0025380.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Xu J, Lodes MJ, Secret H, Benson DR, Meagher MJ, Stolk JA;
PI Wang T, Jiang Y, Smith CL, King GE, Wang A, Clapper JD;
PI Skeiky YAW, Fanger GR, Vedvick TS, Carter D;
XX
DR WPI; 2003-067548/06.
DR N-PSDB; ABZ33691.
XX
XX New polynucleotide, useful for the preparation of a composition for
PT stimulating an immune response against, or treating, cancer -
PT
XX
XX Disclosure; Page 465; 537pp; English.
XX
XX The present invention describes compounds (I) for the immunotherapy and
CC diagnosis of colon cancer. Also described: (1) a method for detecting
CC the presence of cancer in a patient; (2) a method for stimulating and/or
CC expanding T cells specific for a tumour protein; (3) an isolated T cell
CC population comprising T cells prepared by the method of (2); (4) a method
CC for stimulating an immune response in a patient; (5) a method for
CC treating cancer in a patient; and (6) a method for inhibiting the
CC development of cancer in a patient. (I) have immunostimulant and
CC cytostatic activities and can be used in vaccines. ABZ32646 to ABZ33725
CC and ABP55343 to ABP55391 represent human colon cancer/tumour related
CC sequences used in the exemplification of the present invention.
XX
XX Sequence 158 AA;
SQ
Query Match 100.0%; Score 878; DB 24; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.3e-84;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASRSNRLILLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLNWSDAELECO 60
DB 1 MASRSNRLILLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLNWSDAELECO 60
QY 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
DB 61 YNGAHLASILSLKEASTIAEYISGYORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
QY 121 KSMGGNKHCAEMSSNNNFLTWSNCCNKRQHFCLKYRP 158
DB 121 KSMGGNKHCAEMSSNNNFLTWSNCCNKRQHFCLKYRP 158
Search completed: December 31, 2003, 09:13:34
Job time : 72 secs

THIS PAGE BLANK (USPTO)

L3 ANSWER 58 OF 59 BIOSIS COPYRIGHT 2004 BIOLOGICAL ABSTRACTS INC. on STN
DUPLICATE 4
AN 1993:432762 BIOSIS
DN PREV199396087387
TI A **gene homologous** to the **reg** gene is
expressed in the **human** pancreas.
AU Bartoli, Catherine; Gharib, Bouchra; Giorgi, Dominique; Sansonetti, Alice;
Dagorn, Jean-Charles; Berge-LeFranc, Jean-Louis [Reprint author]
CS Unite 315 l'INSERM, 46 Boulevard Gaye, 13258 Marseille Cedex 9, France
SO FEBS (Federation of European Biochemical Societies) Letters, (1993) Vol.
327, No. 3, pp. 289-293.
CODEN: FEBLAL. ISSN: 0014-5793.
DT Article
LA English
OS Genbank-L08010
ED Entered STN: 22 Sep 1993
Last Updated on STN: 6 Nov 1993
AB We have determined the nucleotide sequence of **reg1** a **human**
genomic DNA fragment homologous to the **reg** gene which is
expressed in the exocrine pancreas and regenerating islets.
Sequence comparisons of **reg** and **reg1** suggested similar
exon-intron organisation. Based on this assumption, specific
oligonucleotides for **reg1** exons were used to demonstrate expression of the
reg1 gene in pancreas and liver, The proteins encoded by **reg** and
reg1 comprise 166 amino acids and differ by 22 amino acids only.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 31, 2003, 09:11:17 ; Search time 38 Seconds
(without alignments)
195.532 Million cell updates/sec

Title: US-09-525-041-2

Perfect score: 878

Sequence: 1 MASRMRLLLLLSCLAKTGV.....LTWSSNECKRQHFLCKYRP 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	258.5	29.4	132	1	ACAL_ANSN	P83300	anser anser
2	254.5	29.0	165	1	LITL_MOUSE	P43137	anser musculus
3	254.5	29.0	165	1	LITH_RAT	P10758	rattus norv
4	246.5	28.1	174	1	PAP3_MOUSE	O09049	mus musculus
5	242	27.6	175	1	PAP2_MOUSE	O09037	mus musculus
6	240	27.3	166	1	LITB_HUMAN	P48304	homo sapien
7	236.5	26.9	173	1	LIT2_MOUSE	O08731	mus musculus
8	233	26.5	135	1	LECG_CROAT	P21963	crotalus at
9	231	26.3	174	1	PAP3_RAT	P42854	rattus norv
10	231	26.3	175	1	LITH_BOVIN	P23132	bos taurus
11	228.5	26.0	174	1	PAP2_RAT	P35231	rattus norv
12	225.5	25.7	166	1	LITA_HUMAN	P05451	homo sapien
13	222.5	25.3	172	1	LECA_PLEWA	O02988	pleurodeles
14	221	25.2	175	1	PAP1_HUMAN	O06141	homo sapien
15	218.5	24.9	158	1	LECG_TRIST	O9Y9P1	trimeresuru
16	203	23.1	175	1	PAP1_RAT	P25031	rattus norv
17	202.5	23.1	123	1	ECHE_ECHCA	P81996	echis carin
18	202.5	23.1	152	1	IXA_TRIFL	P23806	trimeresuru
19	199.5	22.7	134	1	ABA2_TRIAB	P81112	trimeresuru
20	197.5	22.5	125	1	ABA3_TRIAB	P81113	trimeresuru
21	196.5	22.4	146	1	IXB_TRIFL	P23807	trimeresuru
22	194	22.0	175	1	PAP1_MOUSE	P35230	mus musculus
23	193	22.0	912	1	PGCB_BOVIN	Q28062	bos taurus
24	191	21.8	883	1	PGCB_MOUSE	O61361	mus musculus
25	190	21.6	175	1	PGCB_HUMAN	Q92778	homo sapien
26	189	21.5	133	1	BOTA_BOTJA	P22029	bothrops ja
27	186.5	21.2	131	1	ABBA_TRIAB	P81111	trimeresuru
28	186.5	21.2	132	1	ABBA_TRIAB	P81115	trimeresuru
29	185.5	21.1	125	1	BOTB_BOTJA	P22030	bothrops ja
30	185	21.1	883	1	PGCB_RAT	P55068	rattus norv
31	185	21.1	1257	1	PGCN_RAT	P55067	rattus norv
32	182.5	20.8	123	1	ABA4_TRIAB	P81114	trimeresuru
33	182.5	20.8	301	1	LECI_MOUSE	P24721	mus musculus

ALIGNMENTS

RESULT 1

ID	ACAL_ANSN	STANDARD;	PRT;	132 AA.
AC	P83300;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Ansocalcin.			
OS	Anser anser (Western graylag goose).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anser.			
OX	NCBI_TaxID=8844;			
RN	[1]			
RP	SEQUENCE, FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, AND MASS			
RP	SPECTROMETRY.			
RC	TISSUE=Eggshell matrix;			
RX	MEDLINE=22439773; PubMed=12431998;			
RA	Lakshminarayanan R., Valiyaveetil S., Rao V.S., Kini R.M.;			
RT	"Purification, characterization, and in vitro mineralization studies			
RL	J. Biol. Chem. 278:2928-2936(2003).			
CC	-!- FUNCTION: Induces spherical aggregates of calcite crystals in			
CC	vitro. Believed to play an active role in the eggshell			
CC	calcification.			
CC	-!- SUBUNIT: Homodimer or homotrimer.			
CC	-!- SUBCELLULAR LOCATION: Eggshell matrix. May be present in the shell			
CC	glands on the walls of oviduct and incorporated into the shell			
CC	structure during its formation.			
CC	-!- MASS SPECTROMETRY: MW=15342; METHOD=Electrospray.			
CC	-!- SIMILARITY: Contains 1 C-type lectin family domain.			
DR	InterPro; IPR002353; AntifreezeII.			
DR	InterPro; IPR001304; Lectin C.			
DR	Pfam; PF00059; Lectin c. 1.			
DR	PRINTS; PR01504; PNCRFATITAP.			
DR	PRINTS; PRO0356; ANTIFREEZEII.			
DR	SMART; SM00034; CLECT; 1.			
DR	PROSITE; PS00615; C-TYPE LECTIN 1; 1.			
DR	PROSITE; PS00041; C-TYPE LECTIN 2; 1.			
KW	Lectin.			
FT	DOMAIN	1	132	C-TYPE LECTIN (LONG FORM) (BY
FT				SIMILARITY).
FT	DISULFID	3	14	BY SIMILARITY.
FT	DISULFID	31	128	BY SIMILARITY.
FT	DISULFID	103	120	BY SIMILARITY.
SQ	SEQUENCE	132 AA;	15347 MW;	36CE42EA4572E6B7 CRC64;

Query Match 29.4%; Score 258.5; DB 1; Length 132;

Best Local Similarity 34.3%; Pred. No. 4.5e-19;

Matches 46; Conservative 25; Mismatches 50; Indels 13; Gaps 3;

QY 30 CAPGPHKSNKYGVFKLRNWSDAELSCQVGNCAHLASILSLKEASTIAIYISGYORS 89

DB 3 CPKGLDPRGSCYGFQGLTWKAEKCKVIHAGCHLASLHSPBEHAAVFIAKFORR 62

QY 90 QP---IWGLHDPQKROQWQ----WIDGAWLYRSWGSKMGNGKCAEMSSNNPLTWS 142


```

CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; L07512; AAA41533.1; -
CC ENBL; M62930; AAA41974.1; -
CC ENBL; M8962; AAA42028.1; -
CC ENBL; D26164; BAA05149.1; -
CC PIR; A28351; A28351.
CC HSSP; P05451; ILIT.
CC InterPro; IPR001304; Lectin C.
CC InterPro; IPR003990; Pancreatins_ac.
CC Pfam; PF00059; lectin C; 1.
CC PRINTS; PR01504; PNCREATITSAP.
CC SMART; SM00034; CLECT; 1.
CC SIGNAL; 1 21 LITHOSTATHINE.
CC CHAIN 22 165 C-TYPE LECTIN (LONG FORM).
CC DOMAIN 33 163 BY SIMILARITY.
CC DISULFID 35 46 BY SIMILARITY.
CC DISULFID 63 161 BY SIMILARITY.
CC DISULFID 136 153 BY SIMILARITY.
CC CARBOHYD 129 129 N-LINKED (GLCNAC... ) (POTENTIAL).
CC SEQUENCE 165 AA; 18672 MW; 9861EB236B82CF8A CRC64;
CC -----
Query Match 29.0%; Score 254.5; DB 1; Length 165;
Best Local Similarity 30.6%; Pred. No. 1.4e-18;
Matches 52; Conservative 38; Mismatches 51; Indels 29; Gaps 7;
QY 7 RLLLSCLAKTGVLDIIMRPS-----CAPGWFYHKSNCYGYFRKLRNW 51
DB 5 KYFILLSCL-----MVLSPSQGQAEEDLPASRITCEGSNAYSYCYFFMEDHLSW 56
QY 52 SDAELSCQVGNCAHLASILSKEASTIAEYI--SGYQRSOPTIWIGLHDPQKQWQWID 109
DB 57 AEADLFQNNNSG-YLVSVLSQEGFLASLIKESGTTAAN-VWIGLHDPQNNRRWHWS 114
QY 110 GAMLYRSW-SGKSMGNK-HCAEMSNNNFLTWSSNECNKQHFCKYR 157
DB 115 GSLFLYKSWDTGTPNNNSGCVSVTSNSGYKKWRDSCDAQSFVCKFK 164
RESULT 4
PAP3_MOUSE STANDARD; PRT; 174 AA.
AC Q09049;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pancreatitis-associated protein 3 precursor (REG III-gamma).
GN PAP3 OR REG3G.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver, and Pancreas;
RX MEDLINE=97208868; PubMed=9055810;
RA Narushima Y., Unno M., Nakagawara K.-I., Mori M., Miyashita H.,
RA Suzuki Y., Noguchi N., Takasawa S., Kumagai T., Yonekura H.,
RA Okamoto H.;
RT "Structure, chromosomal localization and expression of mouse genes
RT encoding type III Reg, RegIII alpha, RegIII beta, RegIII gamma.";
RT Gene 185:159-168(1997).
CC -----
CC -!- FUNCTION: MIGHT BE A STRESS PROTEIN INVOLVED IN THE CONTROL OF
CC BACTERIAL PROLIFERATION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN THE SMALL
CC INTESTINE. MODERATELY IN COLON AND AT AN EXTREMELY LOW LEVEL IN
CC HEALTHY PANCREAS.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; D63361; BAA18930.1; -
CC ENBL; D63362; BAA18931.1; -
CC HSSP; P05451; ILIT.
CC MGD; MGI:109406; Reg3g.
CC InterPro; IPR001304; Lectin C.
CC InterPro; IPR003990; Pancreatins_ac.
CC Pfam; PF00059; lectin C; 1.
CC PRINTS; PR01504; PNCREATITSAP.
CC SMART; SM00034; CLECT; 1.
CC SIGNAL; 1 26 POTENTIAL.
CC CHAIN 27 174 PANCREATITIS-ASSOCIATED PROTEIN 3.
CC DOMAIN 38 172 C-TYPE LECTIN (LONG FORM).
CC DISULFID 40 51 BY SIMILARITY.
CC DISULFID 68 170 BY SIMILARITY.
CC DISULFID 145 162 BY SIMILARITY.
CC SEQUENCE 174 AA; 19307 MW; 5575E9E56A4D8CEP CRC64;
CC -----
Query Match 28.1%; Score 246.5; DB 1; Length 174;
Best Local Similarity 35.2%; Pred. No. 9.8e-18;
Matches 57; Conservative 24; Mismatches 66; Indels 15; Gaps 6;
QY 10 LLLSCL-AKTGVLDII-----MRPSCAPGWFYHKSNCYGYFRKLRNWSDAELSCQV 62
DB 13 MLLSCLMLLSQVQGVAKKQDAPSSRSCPKSGRAYSGYCYALFVSQKNWYDAMACQKRP 72
QY 63 NGAHLASILSKEASTIAEYI--SGYQRSOPTIWIGLHDP-----QKQWQWIDGAMLYR 116
DB 73 SG-HLVSVLSGAEASFLLSMKSSGNSQVYVWIGLHDPDLTGLGYEPNRRGWEWSNADVMNYI 131
QY 117 SW-SGKSMGNK-HCAEMSNNNFLTWSSNECNKQHFCKYR 157
DB 132 NWETNPSSSGNHCGTSLRASGFLKWRNYCNLELPYVCKPK 173
RESULT 5
PAP2_MOUSE STANDARD; PRT; 175 AA.
AC Q09037;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pancreatitis-associated protein 2 precursor (Lithostathine 3) (Islet
DE of langerhans regenerating protein 3) (REG 3) (REG III-alpha).
GN PAP2 OR REG3A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver, and Pancreas;
RX MEDLINE=97208868; PubMed=9055810;
RA Narushima Y., Unno M., Nakagawara K.-I., Mori M., Miyashita H.,
RA Suzuki Y., Noguchi N., Takasawa S., Kumagai T., Yonekura H.,
RA Narushima Y., Unno M., Nakagawara K.-I., Mori M., Miyashita H.,
RA Suzuki Y., Noguchi N., Takasawa S., Kumagai T., Yonekura H.,

```

RA Okamoto H.;
 RT "Structure, chromosomal localization and expression of mouse genes
 RL encoding type III Reg, RegIII alpha, RegIII beta, RegIII gamma.";
 RL Gene 185:159-168(1997).
 CC -!- FUNCTION: MIGHT BE A STRESS PROTEIN INVOLVED IN THE CONTROL OF
 CC BACTERIAL PROLIFERATION.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: SMALL INTESTINE, AND PANCREAS.
 CC -!- DISEASE: OVEREXPRESSED DURING THE ACUTE PHASE OF PANCREATITIS.
 CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D63356; BAA18925.1; -;
 DR EMBL; D63357; BAA18926.1; -;
 DR EMBL; D63358; BAA18927.1; -;
 DR HSSP; P05451; ILIT.
 DR MGD; MGI:109408; Reg3a.
 DR InterPro; IPR001304; Lectin C.
 DR InterPro; IPR003990; Pancreatias_ac.
 DR Pfam; PF00059; lectin c; 1.
 DR PRINTS; PR01504; PNCREATITSAP.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
 DR PROSITE; PS50041; C-TYPE LECTIN 2; 1.
 KW Signal; Lectin; Inflammatory response; Acute phase; Multigene family.
 FT SIGNAL 1 26 BY SIMILARITY.
 FT CHAIN 27 175 PANCREATITIS-ASSOCIATED PROTEIN 2.
 FT DOMAIN 38 173 C-TYPE LECTIN (LONG FORM).
 FT DISULFID 40 51 BY SIMILARITY.
 FT DISULFID 68 171 BY SIMILARITY.
 FT DISULFID 146 163 BY SIMILARITY.
 SQ SEQUENCE 175 AA; 19539 MW; 3E311B3976E80F11 CRC64;
 Query Match 27.6%; Score 242; DB 1; Length 175;
 Best Local Similarity 34.4%; Pred. No. 2.8e-17;
 Matches 56; Conservative 27; Mismatches 64; Indels 16; Gaps 6;
 QY 10 LLSCLAKT-GVLGDIIM-----RPSCAPGWFYHKSNVCYFKLRNWSDAELCOISYG 62
 DB 13 MLLSCLLFVQVQGGDFQKEVPSRTSCPMGYKAYRSHCYALVMTPKSWFQADLVCKRRP 72
 QY 63 NGAHLASLSLKEASTIAEYISG-YQRSQPIWIGLHDPKRRQ-----WOWIDGAMLYLR 116
 DB 73 SG-HLVSLTSGGEASFVSLVNGRVDNYQDIWIGLHDTWCGQPNQGGGWEHNSDVLNLYL 131
 QY 117 SWSG--KSMGNKGKCAEMSSNNFLTWSSNCKNRQHFCKYR 157
 DB 132 NWDGDPSTVNRHGCHGSLTASSGFLKMGDYCYCDGTLPPVCKFK 174
 RESULT 6
 LITB HUMAN
 ID LITB_HUMAN STANDARD; PRT; 166 AA.
 AC P48304;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lithostathine 1 beta precursor (Regenerating protein I beta).
 GN REGIB OR REGI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;

RX MEDLINE=94153997; PubMed=8110835;
 RA Morizumi S., Watanabe T., Unno M., Nakagawara K.I., Suzuki Y.,
 RA Miyashita H., Yonekura H., Okamoto H.;
 RT "Isolation, structural determination and expression of a novel reg
 RL gene, human regi beta.";
 RL Biochim. Biophys. Acta 1217:199-202(1994).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93351647; PubMed=8348956;
 RA Bartoli C., Gharib B., Giorgi D., Sansonetti A., Dagorn J.-C.,
 RA Berge-Lefranc U.;
 RT "A gene homologous to the reg gene is expressed in the human
 RT pancreas.";
 RN FEBS Lett. 327:289-293(1993).
 RN [3]
 RP CARBOHYDRATE-LINKAGE SITE.
 RX MEDLINE=95331286; PubMed=7607222;
 RA De Reggi M., Capon C., Gharib B., Wieruszski J.M., Michel R.,
 RA Fournet B.;
 RA "The glycan moiety of human pancreatic lithostathine. Structure
 RT characterization and possible pathophysiological implications.";
 RL Eur. J. Biochem. 230:503-510(1995).
 CC -!- FUNCTION: MIGHT ACT AS AN INHIBITOR OF SPONTANEOUS CALCIUM
 CC CARBONATE PRECIPITATION. MAY BE ASSOCIATED WITH NEURONAL
 CC SPROUTING IN BRAIN, AND WITH BRAIN AND PANCREAS REGENERATION.
 CC -!- PTM: ALL O-LINKED GLYCANS CONSIST OF GAL-GLCNAC-GAL-GALNAC
 CC TETRASACCHARIDE CORE AND GET ELONGATED (MICROHETEROGENEITY).
 CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D17291; BAA04124.1; -;
 DR EMBL; D16816; BAA04091.1; -;
 DR EMBL; L08010; AAA18204.1; -;
 DR PIR; S34591; RGHUB.
 DR HSSP; P05451; ILIT.
 DR Genew; HGNC:9952; REGIB.
 DR MIM; 167771; -;
 DR GO; GO:0008283; P-cell proliferation; TAS.
 DR InterPro; IPR001304; Lectin C.
 DR InterPro; IPR003990; Pancreatias_ac.
 DR Pfam; PF00059; lectin c; 1.
 DR PRINTS; PR01504; PNCREATITSAP.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
 DR PROSITE; PS50041; C-TYPE LECTIN 2; 1.
 KW Glycoprotein; Signal; Lectin; Pyridolone carboxylic acid.
 FT SIGNAL 1 22 BY SIMILARITY.
 FT CHAIN 23 166 LITHOSTATHINE 1 BETA.
 FT DOMAIN 34 164 C-TYPE LECTIN (LONG FORM).
 FT MOD_RES 23 23 PYRROLIDONE CARBOXYLIC ACID
 FT (BY SIMILARITY).
 FT CARBOHYD 27 27 O-LINKED (GALNAC. .) (MUCIN TYPE).
 FT DISULFID 36 47 BY SIMILARITY.
 FT DISULFID 64 162 BY SIMILARITY.
 FT DISULFID 137 154 BY SIMILARITY.
 SQ SEQUENCE 166 AA; 18665 MW; D1DC20E11AE5DDE8 CRC64;
 Query Match 27.3%; Score 240; DB 1; Length 166;
 Best Local Similarity 34.3%; Pred. No. 4.2e-17;
 Matches 46; Conservative 27; Mismatches 57; Indels 4; Gaps 4;
 QY 27 RPSCAPGWFYHKSNVCYFKLRNWSDAELCOISYGNGAHLASLSLKEASTIAEYISG 86
 DB 33 RISCPEGTNAYRSYCYFYFNPETWVDADLYCONMNSG-NLVSVLTQAGAFVASLIKES 91
 QY 87 QRSQP-IWIGLHDPKRRQWQWIDGAMLYLRSW-SGKSMGGNKH-HCAEMSSNNFLTWSS 143

Query Match	26.9%;	Score 236.5;	DB 1;	Length 173;
Best Local Similarity	28.7%;	Pred. No. 9.8e-17;		

```

Db 62 GOENWIGLRKDKDFSEWTDTRSDCTDYLTDKNQPDHYQNKQKFCVELVSLTGRLWQDQ 121
QY 145 ECKRQHELCK 155
Db 122 VCESKDAFLCQ 132

RESULT 9
PAP3 RAT
ID PAP3 RAT STANDARD; PRT; 174 AA.
AC P42854;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pancreatitis-associated protein 3 precursor.
GN PAP3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Intestine;
RX MEDLINE=94060113; PubMed=8241280;
RA Frigerio J.-M., Dusetti N.J., Garrido P., Dagorn J.C., Iovanna J.L.;
RT "The pancreatitis associated protein III (PAP III), a new member of
RT the PAP gene family.";
RL Biochim. Biophys. Acta 1216:329-331(1993).
CC -!- FUNCTION: MIGHT BE A STRESS PROTEIN INVOLVED IN THE CONTROL OF
CC BACTERIAL PROLIFERATION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; L20869; AAA41809.1; -.
DR EMBL; U09193; AAA79231.1; -.
DR PIR; S54979; S54979.
DR HSP; P05451; 1LIT.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR003990; Pancreatis_ac.
DR Pfam; PF00059; lectin c; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1. LECTIN_1; FALSE NEG.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE NEG.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Signal; Lectin; Inflammatory response; Acute phase.
FT SIGNAL 1 26
FT CHAIN 27 174 PANCREATITIS-ASSOCIATED PROTEIN 3.
FT DONAIN 38 172 C-TYPE LECTIN (LONG FORM).
FT DISULFID 40 51 BY SIMILARITY.
FT DISULFID 68 170 BY SIMILARITY.
FT DISULFID 145 162 BY SIMILARITY.
SQ SEQUENCE 174 AA; 19143 MW; 5CED2E887C46E45C CRC64;

Query Match 26.3%; Score 231; DB 1; Length 174;
Best Local Similarity 32.4%; Fred. No. 3.5e-16;
Matches 55; Conservative 28; Mismatches 73; Indels 14; Gaps 5;

QY 1 MASRSMRLLLLSLCKATGVLGD-----IMRPSCAPGFYHNSCYGVRKRLNNSDA 54
Db 5 VALTTMSWLLSLMLLSQVQGEDAKEDVPTSRISCPKGSRAYGSCYVALPSVSKSNFDA 64
QY 55 ELECQSYNGAHLASILSLKKAETAEYI-SGYQRSDPIWGLHDPKQRQO-----WQVI 108
Db 65 DLACQKPSG-HLVSVLSGSEASFSVSLIKSGNSGNQVWITGLDPTIGQPNRGGWNS 123

```

```

QY 109 DGAMLYRSW-SGKSMGKNKHCAMSSNNNFLTWSNCKNKRQHFCKYR 157
Db 124 NADVMYFNWETNPSSVSGSHCGTLTRASGFLAWRENNCISELPYVCKPK 173

RESULT 10
LITH BOVIN
ID LITH BOVIN STANDARD; PRT; 175 AA.
AC P23132;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lithostathine precursor (Pancreatic stone protein) (PSP) (Pancreatic
DE thread protein) (PTP) (Islet of Langerhans regenerating protein) (REG)
DE (Islet cells regeneration factor) (ICRF).
GN PTP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=90368981; PubMed=2394826;
RA de la Monte S.M., Ozturk M., Wands J.R.;
RT "Enhanced expression of an exocrine pancreatic protein in Alzheimer's
RT disease and the developing human brain.";
RL J. Clin. Invest. 86:1004-1013(1990).
RN [2]
SEQUENCE OF 38-138 AND 141-175.
RX MEDLINE=91197388; PubMed=2085387;
RA Cai L., Harris W.R., Marshak D.R., Gross J., Crabb J.W.;
RT "Structural analysis of bovine pancreatic thread protein.";
RL J. Protein Chem. 9:623-632(1990).
RN [3]
SEQUENCE OF 38-85 AND 141-175.
RX MEDLINE=85298214; PubMed=3862086;
RA Gross J., Brauer A.W., Bringham R.F., Corbett C., Margolies M.N.;
RT "An unusual bovine pancreatic protein exhibiting pH-dependent
RT globule-fibril transformation and unique amino acid sequence.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:5627-5631(1985).
CC -!- FUNCTION: MIGHT ACT AS AN INHIBITOR OF SPONTANEOUS CALCIUM
CC CARBONATE PRECIPITATION.
CC -!- SUBUNIT: CLEAVED TO GIVE AN A CHAIN AND A B CHAIN JOINED BY A
CC DISULFIDE BOND.
CC -!- TISSUE SPECIFICITY: IN PANCREATIC ACINAR CELLS.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M59794; AAA30750.1; -.
DR PIR; A37194; A37194.
DR HSP; P05451; 1QDD.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR003990; Pancreatis_ac.
DR Pfam; PF00059; lectin c; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Glycoprotein; Signal; Lectin.
FT SIGNAL 1 26 POTENTIAL.
FT PROPEP 27 37 POTENTIAL.
FT CHAIN 38 175 LITHOSTATHINE.
FT CHAIN 38 138 A CHAIN.
FT CHAIN 141 175 B CHAIN.

```

FT DOMAIN 38 173 C-TYPE LECTIN (LONG FORM).
 FT DISULFID 40 51 BY SIMILARITY.
 FT DISULFID 68 171 BY SIMILARITY.
 FT DISULFID 146 163 BY SIMILARITY.
 FT CONFLICT 84 85 EE -> FF (IN REF. 3).
 SQ SEQUENCE 175 AA; 19334 MW; C270BE70B7E91D6A CRC64;

Query Match 26.3%; Score 231; DB:1; Length 175;
 Best Local Similarity 30.9%; Pred. No. 3.5e-16;
 Matches 51; Conservative 30; Mismatches 62; Indels 22; Gaps 6;

QY 10 LLLSLAKTGVLDI-----IMRSCAPGWYHKSNCYGYFRKLRNWSDALECO 59
 DB 13 MLLSCLM---LLSQIQGENSKELSPARISCPGSMAYRSHCYALFKPTKTMWDADIACQ 69
 QY 60 SYNGAHLASILSLKEASTIAEYISQYRQSQ-PIWIGLHDPOKROQ-----WQWIDGAMY 113
 DB 70 KPSPG-HLVSVLSGAESEFVASLVNNTQSDIWIHLDPTEGSEANAGGHEWISNDVL 128
 QY 114 LYRSM--SGKSMGKNGKHCAMSSNNFLTWSSNECNKROHFLCKY 156
 DB 129 NYVAMETDPAATSSPGYCGSLRSRSGYLKWRDHNCNLNLPYVCKF 173

RESULT 11
 ID P2P2 RAT STANDARD; PRT; 174 AA.
 AC P35231;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DE Pancreatitis-associated protein 2 precursor (Lithostathine 3) (Islet
 DE of langerhans regenerating protein 3) (REG 3).
 GN PAP2 OR REG3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=93378971; PubMed=8369291;
 RA Frigerio J.-M., Duseti N.J., Keim V., Dagorn J.C., Iovanna J.L.;
 RT "Identification of a second rat pancreatitis-associated protein.
 RT Messenger RNA cloning, gene structure, and expression during acute
 RL pancreatitis.";
 RL Biochemistry 32:9236-9241(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar;
 RX MEDLINE=94314238; PubMed=8039722;
 RA Suzuki Y., Yonekura H., Watanabe T., Unno M., Moriizumi S.,
 RA Miyashita H., Okamoto H.;
 RT "Structure and expression of a novel rat RegIII gene.";
 RL Gene 144:315-316(1994).
 CC -!- FUNCTION: MIGHT BE A STRESS PROTEIN INVOLVED IN THE CONTROL OF
 CC BACTERIAL PROLIFERATION.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: LOW EXPRESSION FOUND IN HEALTHY PANCREAS.
 CC -!- INDUCTION: APPEARS IN PANCREATIC JUICE AFTER INDUCTION OF
 CC PANCREATIC INFLAMMATION.
 CC -!- DISEASE: OVEREXPRESSED DURING THE ACUTE PHASE OF PANCREATITIS.
 CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DB EMBL; LI0230; AAA41808.1; -.

DR EMBL; LI0229; AAA02980.1; -.
 DR EMBL; D26078; BAA05071.1; -.
 DR EMBL; D23676; BAA04904.1; -.
 DR PIR; A48689; A48689.
 DR PIR; I60296; I83377.
 DR HSP; P03451; ILIT.
 DR InterPro; IPR001304; LECTIN_C.
 DR InterPro; IPR003990; Pancreatins_ac.
 DR Pfam; PF00059; lectin C; 1.
 DR PRINTS; PR01504; PNCREATITSAP.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
 DR PROSITE; PS00041; C-TYPE_LECTIN_2; 1.
 KW Signal; Lectin; Inflammatory response; Acute phase.
 FT SIGNAL 1 25
 FT CHAIN 26 174 PANCREATITIS-ASSOCIATED PROTEIN 2.
 FT DOMAIN 37 172 C-TYPE LECTIN (LONG FORM).
 FT DISULFID 39 50 BY SIMILARITY.
 FT DISULFID 67 170 BY SIMILARITY.
 FT DISULFID 145 162 BY SIMILARITY.
 FT CONFLICT 104 104 W -> G (IN REF. 2).
 SQ SEQUENCE 174 AA; 19599 MW; 4530B884496E5888 CRC64;

Query Match 26.0%; Score 228.5; DB:1; Length 174;
 Best Local Similarity 34.5%; Pred. No. 6.3e-16;
 Matches 48; Conservative 25; Mismatches 57; Indels 9; Gaps 4;

QY 27 RPSCAPGWYHKSNCYGYFRKLRNWSDALESCQSYNGAHLASILSLKEASTIAEYISG- 85
 DB 36 RTSCPMGSKAYRSYCYTLVTTLSKWFQADLACQKRPQG-HLVSVLSGGEAFVSVSLVTGR 94
 QY 86 YQRSQPIWIGLHDPOKROQ-----WQWIDGAMYLYRWSG--KSMGKNGKHCAMSSNNNF 138
 DB 95 VNNQDIWIHLDPTEGQPPNGGGEWSNSDVLNVLNWDGDPSSVTVNRGCGSLTATSEF 154
 QY 139 LTWSSNECNKROHFLCKYR 157
 DB 155 LKWDHHDVLPVCKFK 173

RESULT 12
 ID LITA HUMAN STANDARD; PRT; 166 AA.
 AC P05451; P11379;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Lithostathine 1 alpha precursor (pancreatic stone protein) (PSP)
 DE (pancreatic thread protein) (PTP) (Islet of langerhans regenerating
 DE protein) (REG) (Regenerating protein I alpha) (Islet cells
 DE regeneration factor) (ICRP).
 GN REGIA OR REG OR PSPS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88115343; PubMed=2963000;
 RA Terazono K., Yamamoto H., Takasawa S., Shiga K., Yonemura Y.,
 RA Tochino Y., Okamoto H.;
 RT "A novel gene activated in regenerating islets.";
 RL J. Biol. Chem. 263:2111-2114(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90237042; PubMed=2332435;
 RA Watanabe T., Yonekura H., Terazono K., Yamamoto H., Okamoto H.;
 RT "Complete nucleotide sequence of human reg gene and its expression in
 RT normal and tumoral tissues. The reg protein, pancreatic stone
 RT protein, and pancreatic thread protein are one and the same product
 RT of the gene.";
 RL J. Biol. Chem. 265:7432-7439(1990).
 RN [3]

RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=89292148; PubMed=2525567;
RA Giorgi D., Bernard J.-P., Rouquier S., Iovanna J., Sarles H.,
RN Dagorn J.-C.;
RT "Secretory pancreatic stone protein messenger RNA. Nucleotide
sequence and expression in chronic calcifying pancreatitis";
RL J. Clin. Invest. 84:100-106(1989).
RN [4]
RP SEQUENCE FROM N.A.
RA Boonyaisawat W., Tandhanand-Banchuin N., Vannaseng S.,
RA Yenchitsomanus P.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP SEQUENCE OF 34-166.
RX MEDLINE=88029417; PubMed=3665916;
RA de Caro A.M., Bonicel J.J., Rouimi P., de Caro J.D., Sarles H.,
RA Roverly M.;
RT "Complete amino acid sequence of an immunoreactive form of human
pancreatic stone protein isolated from pancreatic juice";
RL Eur. J. Biochem. 168:201-207(1987).
RN [7]
RP SEQUENCE OF 34-98.
RX MEDLINE=87099950; PubMed=3541906;
RA Montalto G., Bonicel J.J., Multigner L., Roverly M., Sarles H.,
RA de Caro A.M.;
RT "Partial amino acid sequence of human pancreatic stone protein, a
novel pancreatic secretory protein";
RL Biochem. J. 238:227-232(1986).
RN [8]
RP SEQUENCE OF 34-78.
RX MEDLINE=86086356; PubMed=3908481;
RA Gross J., Carlson R.I., Brauer A.W., Margolies M.N., Warshaw A.L.,
RA Wands J.R.;
RT "Isolation, characterization, and distribution of an unusual
pancreatic human secretory protein";
RL J. Clin. Invest. 76:2115-2125(1985).
RN [9]
RP SEQUENCE OF 23-47.
RX MEDLINE=89150292; PubMed=2493268;
RA de Caro A.M., Adrich Z., Fournet B., Capon C., Bonicel J.J.,
RA de Caro J.D., Roverly M.;
RT "N-terminal sequence extension in the glycosylated forms of human
pancreatic stone protein. The 5-oxoproline N-terminal chain is O-
glycosylated on the 5th amino acid residue";
RL Biochim. Biophys. Acta 994:281-284(1989).
RN [10]
RP SEQUENCE OF 33-58.
RX MEDLINE=87219142; PubMed=3108036;
RA de Caro A.M., Bonicel J.J., Rouimi P., de Caro J.D., Sarles H.,
RA Roverly M.;
RT "Complete amino acid sequence of an immunoreactive form of human
pancreatic stone protein isolated from pancreatic juice";
RL Eur. J. Biochem. 168:201-207(1987).
RN [11]
RP SEQUENCE OF 216:195-199(1987).
RX MEDLINE=89350859; PubMed=2764894;
RA Stewart T.A.;
RT "The human reg gene encodes pancreatic stone protein";
RL Biochem. J. 260:622-623(1989).
RN [12]
RP DISULFIDE BONDS.
RX MEDLINE=91032149; PubMed=22626837;
RA Itoh T., Tsuzuki H., Katoh T., Teraoka H., Matsumoto K., Yoshida N.,
RA Terazono K., Watanabe T., Yonekura H., Yamamoto H., Okamoto H.;
RT "Isolation and characterization of human reg protein produced in
Saccharomyces cerevisiae";
RL FEBS Lett. 272:85-88(1990).
RN [13]
RP ALZHEIMER'S DISEASE AND DEVELOPMENTAL EXPRESSION REGULATION.
RX MEDLINE=90368981; PubMed=2394826;
RA de la Monte S.M., Ozturk M., Wands J.R.;
RT "Enhanced expression of an exocrine pancreatic protein in Alzheimer's
disease and the developing human brain";
RL J. Clin. Invest. 86:1004-1013(1990).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE=96256285; PubMed=8654365;
RA Bertrand J.A., Pignol D., Bernard J.-P., Verdier J.-M., Dagorn J.-C.,
RA Fontecilla-Camps J.C.;
RT "Crystal structure of human lithostathine, the pancreatic inhibitor
of stone formation";
RL EMBO J. 15:2678-2684(1996).
RN [15]
RP X-RAY CRYSTALLOGRAPHY (1.30 ANGSTROMS) OF 23-166.
RX MEDLINE=20092874; PubMed=10625646;
RA Gerbaud V., Pignol D., Loret E., Bertrand J.A., Berland Y.,
RA Fontecilla-Camps J.C., Canselier J.P., Gabas N., Verdier J.M.;
RT "Mechanism of calcite crystal growth inhibition by the N-terminal
undecapeptide of lithostathine";
RL J. Biol. Chem. 275:1057-1064(2000).
RN [16]
RP STRUCTURE BY NMR OF 34-164.
RX MEDLINE=97120677; PubMed=8961348;
RA Patard L., Stoven V., Gharib B., Bontems F., Lallemand J.-Y.,
RA de Reggi M.;
RT "What function for human lithostathine? structural investigations by
three-dimensional structure modeling and high-resolution NMR
spectroscopy";
RL Protein Eng. 9:949-957(1996).
CC -1- FUNCTION: MIGHT ACT AS AN INHIBITOR OF SPONTANEOUS CALCIUM
CARBONATE PRECIPITATION. MAY BE ASSOCIATED WITH NEURONAL
SPROUTING IN BRAIN, AND WITH BRAIN AND PANCREAS REGENERATION.
CC -1- TISSUE SPECIFICITY: IN PANCREATIC ACINAR CELLS AND, IN LOWER
LEVELS, IN BRAIN.
CC -1- DEVELOPMENTAL STAGE: HIGH EXPRESSION LEVELS IN FETAL AND INFANT
BRAINS; MUCH LOWER IN ADULT BRAINS.
CC -1- DISEASE: ALZHEIMER'S DISEASE AND DOWN'S SYNDROME PATIENTS SHOW
ENHANCED EXPRESSION OF PSP-RELATED TRANSCRIPTS AND INTRANEUROAL
ACCUMULATION OF PSP-LIKE PROTEINS IN THEIR BRAINS.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M27190; AAA60546.1; -;
DR EMBL; M27189; AAA60545.1; -;
DR EMBL; M18963; AAA36558.1; -;


```

CC -I- SIMILARITY: Contains 1 C-type lectin family domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X69062; CAA48800.1; -.
CC DR PIR: S32489; S32489.
CC DR HSPG: P20693; IHLJ.
CC DR InterPro: IPR001304; Lectin_C.
CC DR InterPro: IPR003990; Pancreatis_ac.
CC DR Pfam: PF00059; lectin_c.1.
CC DR PRINTS: PR01504; PNCREATITSAP.
CC DR SMART: SM00034; CLECT.1.
CC DR PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
CC DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
CC KW Lectin; Signal; Glycoprotein.
CC FT SIGNAL 1 20
CC FT CHAIN 21 172 LECTIN
CC FT DOMAIN 36 172 C-TYPE LECTIN (LONG FORM).
CC FT CARBOHYD 93 93 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT DISULFID 65 170 BY SIMILARITY.
CC FT DISULFID 156 164 BY SIMILARITY.
CC SQ SEQUENCE 172 AA; 20141 MW; 28BB9FF12C136EA3 CRC64;
CC -----
Query Match 25.3%; Score 222.5; DB 1; Length 172;
Best Local Similarity 34.3%; Pred. No. 2.5e-15;
Matches 46; Conservative 22; Mismatches 55; Indels 11; Gaps 4;
CC -----
QY 30 CAPGWFYHKSNCYGVYPRKLRNWSDALEQSYGNGAHLASILSLKEASTIAEYISGVQRS 89
DB 37 CTGPGWCHFNYSYKYP1PNAKSWTDABFYCKLYPGAHLASIHSEDENDFTETFKNSN 96
QY 90 QP-ITWIGLHDPPKQKQWOWIDGAMLYL---RSWSGKSMGNGKHCAEMSSNNFLT---WS 142
DB 97 YPVVWVGSDCYKDRSFVWTDGSDQWYQKWRQWEPNPTGGRPCIDP---NPVTPGLWN 152
QY 143 SNECNKRQHPLCKY 156
DB 153 DEHCDQKQFPFICKY 166
CC -----
RESULT 14
PAP1_HUMAN STANDARD; PRT; 175 AA.
ID ID PAP1_HUMAN
AC Q06141;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DS Pancreatitis-associated protein 1 precursor.
GN PAP OR PAP1 OR HIP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
CX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Pancreas;
RC MEDLINE=93176807; PubMed=7679928;
RA Itoh T., Teraoka H.;
RT "Cloning and tissue-specific expression of cDNAs for the human and
RL mouse homologues of rat pancreatitis-associated protein (PAP).";
RL Biochim. Biophys. Acta 1172:184-186(1993).
RX [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Pancreas;
RC MEDLINE=93107309; PubMed=1469087;
RA Orelli B., Keim V., Masciotra L., Dagorn J.-C., Iovanna J.-L.;
RT "Human pancreatitis-associated protein. Messenger RNA cloning and

```

EMBL; J05412; AAA36559.1; --
EMBL; AF172331; AAD51330.1; --
EMBL; BC005350; RAH05350.1; --
PIR; A35197; RGHUIA.
PIR; A45751; A45751.
PDB; 1LIT; 11-JAN-97.
PDB; 1QDD; 24-JAN-01.
Genew; HGNC:9951; REGIA.
MIM; 167770; --
MIM; 167800; --
GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
InterPro; IPR002353; AntifreezeII.
InterPro; IPR001304; Lectin_C.
InterPro; IPR003990; Pancreatias_ac.
Pfam; PF00059; lectin c; 1.
PRINTS; PR01504; PNCREATITAP.
PRINTS; PR00356; ANTIFREEZEII.
SMART; SM00034; CLECT; 1.
PROSITE; PS00615; C-TYPE LECTIN_1; 1.
PROSITE; PS00041; C-TYPE LECTIN_2; 1.
KW Glycoprotein; Signal; Alzheimer's disease; Down's syndrome; Lectin;
KW 3D-structure; Pyroglutamate carboxylic acid.
FT SIGNAL 1 22

Query Match 25.7%; Score 225.5; DB 1; Length 166;
Best Local Similarity 31.5%; Pred. No. 1.2e-15;
Matches 51; Conservative 30; Mismatches 62; Indels 19; Gaps 6;

QY 10 LLLSL-----AKTVGLDIIIRPSCAPGWYHKNCYCYFVKLRNWSDALECO 59
DB 9 MLICMLFSLSQSQGEATEIPQ---ARISCEPTNAYRSYCYFVFNEDRETWDADLYCQ 65
QY 60 SYGNGAHLASILSLKEASTIAEYI--SGYQRSQPIWGLHDPKQKQOWIDGAMLYRS 117
DB 66 NNMSG-NLVSVLTOEAGAFVSLIKESGTDFFN-VWGLHDPKQKRWHRHSSGSLVYSKS 123
QY 118 W--SKSMGNGKHCHEMSNNFLTWSSNECNKQHFLECKYR 157
DB 124 WGIAPSSVNGYCVSLTSSGTGFKWQKDVPCEDKFSFVCKFK 165

RESULT 13
LECA_PLEWA
ID LECA_PLEWA STANDARD; PRT; 172 AA.
AC Q02988;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lectin precursor.
GN LEC.
OS Pleurodeles waltlii (Iberian ribbed newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
OC Pleurodeles.
NCBI_Taxid=8319;
RN [1]
R1 Tissue=Oviduct;
R2 Tissue=Oviduct;
R3 Tissue=Oviduct;
R4 Tissue=Oviduct;
R5 Tissue=Oviduct;
R6 Tissue=Oviduct;
R7 Tissue=Oviduct;
R8 Tissue=Oviduct;
R9 Tissue=Oviduct;
R10 Tissue=Oviduct;
R11 Tissue=Oviduct;
R12 Tissue=Oviduct;
R13 Tissue=Oviduct;
R14 Tissue=Oviduct;
R15 Tissue=Oviduct;
R16 Tissue=Oviduct;
R17 Tissue=Oviduct;
R18 Tissue=Oviduct;
R19 Tissue=Oviduct;
R20 Tissue=Oviduct;
R21 Tissue=Oviduct;
R22 Tissue=Oviduct;
R23 Tissue=Oviduct;
R24 Tissue=Oviduct;
R25 Tissue=Oviduct;
R26 Tissue=Oviduct;
R27 Tissue=Oviduct;
R28 Tissue=Oviduct;
R29 Tissue=Oviduct;
R30 Tissue=Oviduct;
R31 Tissue=Oviduct;
R32 Tissue=Oviduct;
R33 Tissue=Oviduct;
R34 Tissue=Oviduct;
R35 Tissue=Oviduct;
R36 Tissue=Oviduct;
R37 Tissue=Oviduct;
R38 Tissue=Oviduct;
R39 Tissue=Oviduct;
R40 Tissue=Oviduct;
R41 Tissue=Oviduct;
R42 Tissue=Oviduct;
R43 Tissue=Oviduct;
R44 Tissue=Oviduct;
R45 Tissue=Oviduct;
R46 Tissue=Oviduct;
R47 Tissue=Oviduct;
R48 Tissue=Oviduct;
R49 Tissue=Oviduct;
R50 Tissue=Oviduct;
R51 Tissue=Oviduct;
R52 Tissue=Oviduct;
R53 Tissue=Oviduct;
R54 Tissue=Oviduct;
R55 Tissue=Oviduct;
R56 Tissue=Oviduct;
R57 Tissue=Oviduct;
R58 Tissue=Oviduct;
R59 Tissue=Oviduct;
R60 Tissue=Oviduct;
R61 Tissue=Oviduct;
R62 Tissue=Oviduct;
R63 Tissue=Oviduct;
R64 Tissue=Oviduct;
R65 Tissue=Oviduct;
R66 Tissue=Oviduct;
R67 Tissue=Oviduct;
R68 Tissue=Oviduct;
R69 Tissue=Oviduct;
R70 Tissue=Oviduct;
R71 Tissue=Oviduct;
R72 Tissue=Oviduct;
R73 Tissue=Oviduct;
R74 Tissue=Oviduct;
R75 Tissue=Oviduct;
R76 Tissue=Oviduct;
R77 Tissue=Oviduct;
R78 Tissue=Oviduct;
R79 Tissue=Oviduct;
R80 Tissue=Oviduct;
R81 Tissue=Oviduct;
R82 Tissue=Oviduct;
R83 Tissue=Oviduct;
R84 Tissue=Oviduct;
R85 Tissue=Oviduct;
R86 Tissue=Oviduct;
R87 Tissue=Oviduct;
R88 Tissue=Oviduct;
R89 Tissue=Oviduct;
R90 Tissue=Oviduct;
R91 Tissue=Oviduct;
R92 Tissue=Oviduct;
R93 Tissue=Oviduct;
R94 Tissue=Oviduct;
R95 Tissue=Oviduct;
R96 Tissue=Oviduct;
R97 Tissue=Oviduct;
R98 Tissue=Oviduct;
R99 Tissue=Oviduct;
R100 Tissue=Oviduct;
R101 Tissue=Oviduct;
R102 Tissue=Oviduct;
R103 Tissue=Oviduct;
R104 Tissue=Oviduct;
R105 Tissue=Oviduct;
R106 Tissue=Oviduct;
R107 Tissue=Oviduct;
R108 Tissue=Oviduct;
R109 Tissue=Oviduct;
R110 Tissue=Oviduct;
R111 Tissue=Oviduct;
R112 Tissue=Oviduct;
R113 Tissue=Oviduct;
R114 Tissue=Oviduct;
R115 Tissue=Oviduct;
R116 Tissue=Oviduct;
R117 Tissue=Oviduct;
R118 Tissue=Oviduct;
R119 Tissue=Oviduct;
R120 Tissue=Oviduct;
R121 Tissue=Oviduct;
R122 Tissue=Oviduct;
R123 Tissue=Oviduct;
R124 Tissue=Oviduct;
R125 Tissue=Oviduct;
R126 Tissue=Oviduct;
R127 Tissue=Oviduct;
R128 Tissue=Oviduct;
R129 Tissue=Oviduct;
R130 Tissue=Oviduct;
R131 Tissue=Oviduct;
R132 Tissue=Oviduct;
R133 Tissue=Oviduct;
R134 Tissue=Oviduct;
R135 Tissue=Oviduct;
R136 Tissue=Oviduct;
R137 Tissue=Oviduct;
R138 Tissue=Oviduct;
R139 Tissue=Oviduct;
R140 Tissue=Oviduct;
R141 Tissue=Oviduct;
R142 Tissue=Oviduct;
R143 Tissue=Oviduct;
R144 Tissue=Oviduct;
R145 Tissue=Oviduct;
R146 Tissue=Oviduct;
R147 Tissue=Oviduct;
R148 Tissue=Oviduct;
R149 Tissue=Oviduct;
R150 Tissue=Oviduct;
R151 Tissue=Oviduct;
R152 Tissue=Oviduct;
R153 Tissue=Oviduct;
R154 Tissue=Oviduct;
R155 Tissue=Oviduct;
R156 Tissue=Oviduct;
R157 Tissue=Oviduct;
R158 Tissue=Oviduct;
R159 Tissue=Oviduct;
R160 Tissue=Oviduct;
R161 Tissue=Oviduct;
R162 Tissue=Oviduct;
R163 Tissue=Oviduct;
R164 Tissue=Oviduct;
R165 Tissue=Oviduct;
R166 Tissue=Oviduct;
R167 Tissue=Oviduct;
R168 Tissue=Oviduct;
R169 Tissue=Oviduct;
R170 Tissue=Oviduct;
R171 Tissue=Oviduct;
R172 Tissue=Oviduct;
R173 Tissue=Oviduct;
R174 Tissue=Oviduct;
R175 Tissue=Oviduct;
R176 Tissue=Oviduct;
R177 Tissue=Oviduct;
R178 Tissue=Oviduct;
R179 Tissue=Oviduct;
R180 Tissue=Oviduct;
R181 Tissue=Oviduct;
R182 Tissue=Oviduct;
R183 Tissue=Oviduct;
R184 Tissue=Oviduct;
R185 Tissue=Oviduct;
R186 Tissue=Oviduct;
R187 Tissue=Oviduct;
R188 Tissue=Oviduct;
R189 Tissue=Oviduct;
R190 Tissue=Oviduct;
R191 Tissue=Oviduct;
R192 Tissue=Oviduct;
R193 Tissue=Oviduct;
R194 Tissue=Oviduct;
R195 Tissue=Oviduct;
R196 Tissue=Oviduct;
R197 Tissue=Oviduct;
R198 Tissue=Oviduct;
R199 Tissue=Oviduct;
R200 Tissue=Oviduct;
R201 Tissue=Oviduct;
R202 Tissue=Oviduct;
R203 Tissue=Oviduct;
R204 Tissue=Oviduct;
R205 Tissue=Oviduct;
R206 Tissue=Oviduct;
R207 Tissue=Oviduct;
R208 Tissue=Oviduct;
R209 Tissue=Oviduct;
R210 Tissue=Oviduct;
R211 Tissue=Oviduct;
R212 Tissue=Oviduct;
R213 Tissue=Oviduct;
R214 Tissue=Oviduct;
R215 Tissue=Oviduct;
R216 Tissue=Oviduct;
R217 Tissue=Oviduct;
R218 Tissue=Oviduct;
R219 Tissue=Oviduct;
R220 Tissue=Oviduct;
R221 Tissue=Oviduct;
R222 Tissue=Oviduct;
R223 Tissue=Oviduct;
R224 Tissue=Ovid

expression in pancreatic diseases.";
 RL J. Clin. Invest. 90:2284-2291 (1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=92386513; PubMed=1325291;
 RA Laessle C., Christa L., Simon M.T., Vernier P., Brecht C.;
 RT "A novel gene (HIP) activated in human primary liver cancer.";
 RL Cancer Res. 52:5089-5095 (1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=94245143; PubMed=8188210;
 RA Duseti N.J., Frigerio J.M., Fox M.F., Swallow D.M., Dagorn J.C.,
 RA Iovanna J.L.;
 RT Molecular cloning, genomic organization, and chromosomal
 RT localization of the human pancreatitis-associated protein (PAP)
 RT gene.";
 RL Genomics 19:108-114 (1994).
 CC -!- FUNCTION: MIGHT BE A STRESS PROTEIN INVOLVED IN THE CONTROL OF
 CC BACTERIAL PROLIFERATION.
 CC -!- SUBCELLULAR LOCATION: SECRETED. FOUND IN THE APICAL REGION OF
 CC PANCREATIC ACINAR CELLS.
 CC -!- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN INTESTINE. LOW
 CC EXPRESSION IS FOUND IN HEALTHY PANCREAS.
 CC -!- INDUCTION: APPEARS IN PANCREATIC JUICE AFTER INDUCTION OF
 CC PANCREATIC INFLAMMATION.
 CC -!- DISEASE: OVEREXPRESSED DURING THE ACUTE PHASE OF PANCREATITIS AND
 CC IN SOME PATIENTS WITH CHRONIC PANCREATITIS.
 CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D13510; BAA02728.1; -;
 DR EMBL; M84337; AAA36415.1; -;
 DR EMBL; S51768; AAB24642.1; -;
 DR EMBL; X68641; CAA48605.1; -;
 DR EMBL; L15533; AAA60020.1; -;
 DR PIR; A49616; A49616.
 DR HSP; P05451; IQDD.
 DR Genew; HGNC:8601; PAP.
 DR MIN; 167805; -;
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0005615; C:extracellular space; TAS.
 DR GO; GO:0005625; C:soluble fraction; TAS.
 DR GO; GO:0005530; F:lectin; TAS.
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR GO; GO:0008283; P:cell proliferation; TAS.
 DR GO; GO:0007275; P:development; TAS.
 DR InterPro; IPR001304; Lectin C.
 DR InterPro; IPR003990; Pancreatias_ac.
 DR Pfam; PF00059; lectin_c; 1.
 DR PRINTS; PR01504; PNCREATIAS.P.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
 DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
 KW Signal; Lectin; Inflammatory response; Acute phase.
 FT SIGNAL 1 26 BY SIMILARITY.
 FT CHAIN 27 175 PANCREATITIS-ASSOCIATED PROTEIN 1.
 FT DOMAIN 38 173 C-TYPE LECTIN (LONG FORM).
 FT DISULFID 40 51 BY SIMILARITY.
 FT DISULFID 68 171 BY SIMILARITY.
 FT DISULFID 146 163 BY SIMILARITY.
 FT CONFLICT 173 175 FTD -> VH (IN REF. 2).
 SQ SEQUENCE 175 AA; 19395 MW; C51149FAC22EB68C CRC64;
 Query Match 25.2%; Score 221; DB 1; Length 175;

Best Local Similarity 32.7%; Pred. No. 3.6e-15;
 Matches 53; Conservative 29; Mismatches 64; Indels 16; Gaps 7;
 QY 10 LLLSCL-AKTGVLDIIMR--PS-----CAPGFYHKSNCYGYFRKLNRNWSDAELCQSYG 62
 DB 13 MLTSLMLLSQVQGEPEQRELFSAIRCPKSGKAYGSHYALFLSPKSTWDADLACQKRP 72
 QY 63 NGAHLASILSLKEASTIABYISGYQSQP-IGWLHDPQKROQ-----QWIDGAMLYLR 116
 DB 73 SG-NLVSVLSGAEQSFVSLVKSIGNSYSVVMVGLHDPQTQTEPNEGGEWSSSDVMNYF 131
 QY 117 SW--SKSMGNGKHCHEMSSNNFLTWSSNECNKROHFLCKY 156
 DB 132 AWERNPTISSPGHCASLSRSTAFLEKWDYNCNVRPLVYCKF 173
 RESULT 15
 LECG TRIST STANDARD; PRT; 158 AA.
 ID LECG TRIST
 AC QSYGPI;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Galactose-binding lectin precursor (tSL).
 OS Trimeresurus stejnegeri (Chinese green tree viper).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Chordata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Crotalinae; Trimeresurus.
 OC NCBI_TaxID=39682;
 RN [1]
 RP SEQUENCE FROM N.A., SEQUENCE OF 24-53, AND MASS SPECTROMETRY.
 RC TISSUE=Venom;
 RX MEDLINE=99348038; PubMed=10417338;
 RA Xu Q., Wu X.-F., Xia Q.-C., Wang K.-Y.;
 RT "Cloning of a galactose-binding lectin from the venom of Trimeresurus
 RT stejnegeri.";
 RL Biochem. J. 341:733-737 (1999).
 RN [2]
 RP SEQUENCE, CARBOHYDRATE-LINKAGE SITE, AND STRUCTURE OF CARBOHYDRATES.
 RC TISSUE=Venom;
 RX PubMed=10561575;
 RA Zeng R., Xu Q., Shao X.-X., Wang K.-Y., Xia Q.-C.;
 RT "Characterization and analysis of a novel glycoprotein from snake
 RT venom using liquid chromatography-electrospray mass spectrometry and
 RT Edman degradation.";
 RL Eur. J. Biochem. 266:352-358 (1999).
 CC -!- FUNCTION: GALACTOSE-BINDING PROTEIN WHICH RECOGNIZES SPECIFIC
 CC CARBOHYDRATE STRUCTURES AND AGGLUTININATE A VARIETY OF ANIMAL CELLS
 CC BY BINDING TO CELL-SURFACE GLYCOPROTEINS AND GLYCOLIPIDS. MAY BE A
 CC CALCIUM-DEPENDENT LECTIN.
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MASS SPECTROMETRY: MW=17924.2; MW_ERR=2.4; METHOD=Electrospray;
 CC RANGE=24-158.
 CC -!- MISCELLANEOUS: Met-33 has been shown to be oxidized to methionine
 CC sulfoxide (Ref.2) but this probably results from sample treatment
 CC prior to mass spectrometry.
 CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF119097; AAD17252.1; -;
 DR HSP; P22897; LEGG.
 DR GlycoSuiteDB; Q9YGP1; -;
 DR InterPro; IPR001304; Lectin_C.
 DR InterPro; IPR003990; Pancreatias_ac.
 DR Pfam; PF00059; lectin_c; 1.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 31, 2003, 09:23:18 ; Search time 75 Seconds
(without alignments)
929,847 Million cell updates/sec

Title: US-09-525-041-2

Perfect score: 878

Sequence: 1 MASRSMRLLLLSLAKTV.....LTWSSNECKRQHFLCKYRP 158

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlpl

-MODEL=frame+ p2n model -DEV=xlpl
-Q/cgn2_1/USPTO_spool_p/US09525041/runat 31122003 091150 11891/app query.fasta 1.327
-DB=issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09525041 @CGN 1 1 103 @runat 31122003 091150 11891 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -XHEADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCITUS COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	878	100.0	614	2	US-08-729-103-2
2	878	100.0	1114	2	US-08-468-413-1
3	878	100.0	1114	3	US-09-162-508-1
4	878	100.0	1114	5	PCT-US95-07169-1
5	224	25.5	777	3	US-09-146-969-1
6	221	25.2	797	2	US-08-464-637-1
7	221	25.2	797	2	US-08-822-261-5
8	221	25.2	797	4	US-09-226-852-5
9	218	24.8	522	1	US-07-778-156-4
10	218	24.8	522	2	US-08-422-166-4
11	218	24.8	798	1	US-07-778-156-12
12	218	24.8	798	2	US-08-822-261-6

13	218	24.8	798	2	US-08-422-166-12	Sequence 12, Appl
14	218	24.8	798	3	US-09-146-969-2	Sequence 2, Appl
15	218	24.8	798	4	US-09-226-852-6	Sequence 6, Appl
16	212	24.1	590	2	US-08-454-557C-32	Sequence 32, Appl
17	212	24.1	590	2	US-08-340-436D-32	Sequence 32, Appl
18	212	24.1	590	2	US-08-450-673C-32	Sequence 32, Appl
19	212	24.1	590	5	PCT-US95-17111A-32	Sequence 32, Appl
20	204	23.2	762	2	US-08-822-261-2	Sequence 2, Appl
21	204	23.2	762	4	US-09-226-852-2	Sequence 2, Appl
22	193	22.0	3259	5	PCT-US95-03747-1	Sequence 1, Appl
23	190	21.6	586	3	US-09-146-969-3	Sequence 3, Appl
24	190	21.6	747	2	US-08-401-530A-1	Sequence 1, Appl
25	190	21.6	747	2	US-08-709-662-1	Sequence 1, Appl
26	190	21.6	793	1	US-07-778-156-1	Sequence 1, Appl
27	190	21.6	793	2	US-08-422-166-1	Sequence 1, Appl
28	189	21.5	1370	3	US-09-111-470-9	Sequence 9, Appl
29	189	21.5	1458	3	US-09-111-470-3	Sequence 3, Appl
30	186.5	21.2	558	1	US-08-909-725-4	Sequence 4, Appl
31	185.5	21.1	474	1	US-07-778-156-8	Sequence 8, Appl
32	185.5	21.1	474	2	US-08-422-166-8	Sequence 8, Appl
33	185	21.1	5191	1	US-08-340-428B-1	Sequence 1, Appl
34	185	21.1	5191	5	PCT-US93-07306-1	Sequence 1, Appl
35	181	20.6	1277	4	US-09-016-434-1186	Sequence 1186, Ap
36	179	20.4	8224	6	5180808-1	Patent No. 5180808
37	173.5	19.8	544	4	US-09-058-740-1	Sequence 1, Appl
38	172.5	19.6	1212	3	US-09-591-435-11	Sequence 11, Appl
39	171.5	19.5	1212	3	US-09-591-435-10	Sequence 10, Appl
40	170.5	19.4	1212	3	US-09-591-435-9	Sequence 9, Appl
41	170.5	19.4	1312	4	US-09-517-605-1	Sequence 1, Appl
42	168.5	19.2	690	2	US-08-612-840A-7	Sequence 7, Appl
43	161	18.3	5169	4	US-09-194-612A-2	Sequence 2, Appl
44	153.5	17.5	2318	4	US-09-620-312D-733	Sequence 733, App
45	153	17.4	4588	3	US-08-840-062-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-729-103-2
; Sequence 2, Application US/08729103
; Patent No. 5837841
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN REG PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729.103
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0138 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

LENGTH: 614 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
LIBRARY: COLNFET02
CLONE: 1310334
US-08-729-103-2

Alignment Scores:
Pred. No.: 7e-100 Length: 614
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-525-041-2 (1-158) x US-08-729-103-2 (1-614)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB 136 ATGGCTTCCAGAGCATGCGCTCTCTATGTCTGAGCTGCTGGCCAAACAGGAGTC 195
QY 21 LeuGlyAspIlelleMetArgProSerCysAlaProGlyTyrPheTyrHisLysSerAsn 40
DB 196 CTGGGTGATATCATCATGAGACCCAGCTGTCTCTGATGGTGTATACCAAGTCCAAAT 255
QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTyrSerAspAlaGluLeuGluCysGlnSer 60
DB 256 TGTCTATGGTTACTTTCAGGAAGCTGAGGAACCTGGTCTGATGCCAGCTCGAGTGTCTAGTCT 315
QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB 316 TACGGAAACGGAGCCACCTGGCATCTATCTGTAGTTTAAAGGAAGCCAGCACATAGCA 375
QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTyrPheLeuHisAspProGln 100
DB 376 GAGTACATAGTGGCTATCAGAGAAGCCAGCCGATATGGATGGCTTGCACGCCACACAG 435
QY 101 LysArgGlnGlnTyrGlnTyrPheAspGlyAlaMetTyrLeuTyrArgSerTyrSerGly 120
DB 436 AAGAGGCGAGCTGCGAGTGGATTTGATGGGGCCATGTATCTGTACAGATCTCTGTCTGGC 495
QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
DB 496 AAGTCCATGGGTGGGAACAGACACTGTGTGAGATGAGCTCCCAATACAACTTTTAACT 555
QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
DB 556 TGGAGCAGCAACGAATGCACAGCGCCACACTTCTGTGCAAGTACCGACCA 609

RESULT 2

US-08-468-413-1
; Sequence 1, Application US/08468413
; Patent No. 5861494
; GENERAL INFORMATION:
; APPLICANT: LI, ET AL.
; TITLE OF INVENTION: Human Colon Specific Gene
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,413
; FILING DATE: 06 JUN 95
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-447
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1114 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; US-08-468-413-1

Alignment Scores:

Pred. No.: 1.71e-99 Length: 1114
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-525-041-2 (1-158) x US-08-468-413-1 (1-1114)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB 111 ATGGCTTCCAGAGCATGCGCTGTCTCTATGTCTGAGCTGCTGGCCAAACAGGAGTC 170
QY 21 LeuGlyAspIlelleMetArgProSerCysAlaProGlyTyrPheTyrHisLysSerAsn 40
DB 171 CTGGGTGATATCATCATGAGACCCAGCTGTCTCTGGATGGTTTACCAAGTCCAAAT 230
QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTyrSerAspAlaGluLeuGluCysGlnSer 60
DB 231 TGTCTATGGTTACTTTCAGGAAGCTGAGGAACCTGGTCTGTATGCCAGCTCGAGTGTCTAGTCT 290
QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB 291 TACGGAAACGGAGCCACCTGGCATCTATCTGTAGTTTAAAGGAAGCCAGCACATAGCA 350
QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTyrPheLeuHisAspProGln 100
DB 351 GAGTACATAGTGGCTATCAGAGAAGCCAGCCGATATGGATGGCTTGCACGCCACACAG 410
QY 101 LysArgGlnGlnTyrGlnTyrPheAspGlyAlaMetTyrLeuTyrArgSerTyrSerGly 120
DB 411 AAGAGGCGAGCTGCGAGTGGATTTGATGGGGCCATGTATCTGTACAGATCTCTGTCTGGC 470
QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
DB 471 AAGTCCATGGGTGGGAACAGACACTGTGTGAGATGAGCTCCCAATACAACTTTTAACT 530
QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
DB 531 TGGAGCAGCAACGAATGCACAGCGCCACACTTCTGTGCAAGTACCGACCA 584

RESULT 3

US-09-162-508-1
; Sequence 1, Application US/09162508
; Patent No. 6080722
; GENERAL INFORMATION:
; APPLICANT: LI, ET AL.
; TITLE OF INVENTION: Human Colon Specific Gene
; NUMBER OF SEQUENCES: 6

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
;; ADDRESSEE: CECCHI, STEWART & OLSTEIN
;; STREET: 6 BECKER FARM ROAD
;; CITY: ROSELAND
;; STATE: NEW JERSEY
;; COUNTRY: USA
;; ZIP: 07068
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5 INCH DISKETTE
;; COMPUTER: IBM PS/2
;; OPERATING SYSTEM: MS-DOS
;; SOFTWARE: WORD PERFECT 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/162,508
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/468,413
;; FILING DATE: 06 JUN 95
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FERRARO, GREGORY D.
;; REGISTRATION NUMBER: 36,134
;; REFERENCE/DOCKET NUMBER: 325800-447
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 201-994-1740
;; TELEFAX: 201-994-1744
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1114 BASE PAIRS
;; TYPE: NUCLEIC ACID
;; STRANDEDNESS: SINGLE
;; TOPOLOGY: LINEAR
;; MOLECULE TYPE: CDNA
US-09-162-508-1

Alignment Scores:
Pred. No.: 1114
Score: 878.00
Matches: 158
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatches: 0
Query Match: 100.00%
Indels: 0
DB: 3
Gaps: 0

US-09-525-041-2 (1-158) x US-09-162-508-1 (1-1114)

QY 1 MetAlaSerArgSerMetArgLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB 111 ATGGCTTCAGAGCATGCGCTGCTCTATTGCTGAGCTGCTGCCCAAAACAGGAGTC 170
QY 21 LeuGlyAspIleleMetArgProSerCysAlaProGlyTTPheTyrHisLysSerAsn 40
DB 171 CTGGGTGATATCATCATGAGACCCAGCTGTCTCTGGATGGTTTTTACCACAGTCCAT 230
QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTTPSerAspAlaGluLeuGluCysGlnSer 60
DB 231 TGCTATGGTTACTTTCAGGAAGCTGAGCACTGGTCTGATGCCGAGCTCGAGTGTCACT 290
QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB 291 TACGGAACCGAGCCCACTGCTGATGAGTGGCCATCTATCTGATGAGTGGTGTCTGCT 350
QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTTPileGlyLeuHisAspProGln 100
DB 351 GAGTACATAGTGGCTATCAGAGAACCCAGCCGATATGATGGCTTGGCTGACGACCCACAG 410
QY 101 LysArgGlnGlnTTPGlnTTPileAspGlyAlaMetTyrLysTyrArgSerTTPSerGly 120
DB 411 AAGAGGCAGCAGTGGCAGTGGATTGATGGGCCATCTATCTGATGAGTGGTGTCTGCTG 470
QY 121 LysSerMetGlyCysAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
DB 471 AAGTCCATGGTGGGAACAAGCACTGTGCTGATGAGTGGCTCCCAATTAACAACCTTTTAACT 530

QY 141 TTPSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
DB 531 TGGAGCAGCAAGCAATGCAACAAGGCCAACACTTCTCTGTGCAAGTACCGACCA 584
RESULT 4
PCT-US95-07169-1
; Sequence 1, Application PC/TUS9507169
; GENERAL INFORMATION:
; APPLICANT: LI, ET AL.
; TITLE OF INVENTION: Human Colon Specific Gene
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07169
; FILING DATE: 06 JUN 95
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-389
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1114 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
PCT-US95-07169-1

Alignment Scores:
Pred. No.: 1114
Score: 878.00
Matches: 158
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatches: 0
Query Match: 100.00%
Indels: 0
DB: 5
Gaps: 0

US-09-525-041-2 (1-158) x PCT-US95-07169-1 (1-1114)

QY 1 MetAlaSerArgSerMetArgLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB 111 ATGGCTTCAGAGCATGCGCTGCTCTATTGCTGAGCTGCTGCCCAAAACAGGAGTC 170
QY 21 LeuGlyAspIleleMetArgProSerCysAlaProGlyTTPheTyrHisLysSerAsn 40
DB 171 CTGGGTGATATCATCATGAGACCCAGCTGTCTCTGGATGGTTTTTACCACAGTCCAT 230
QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTTPSerAspAlaGluLeuGluCysGlnSer 60
DB 231 TGCTATGGTTACTTTCAGGAAGCTGAGCACTGGTCTGATGCCGAGCTCGAGTGTCACT 290
QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB 291 TACGGAACCGAGCCCACTGCTGATGAGTGGCTCCCAATTAACAACCTTTTAACT 350

Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrrpIleGlyLeuHisAspProGln 100
Db 351 GAGTACATAAGTGGCTATCAGAGAGCCAGCGATATGAGTTGGCTTCAGACCCACAG 410
Qy 101 LysArgGlnGlnTrrpGlnTrrpIleAspGlyAlaMetTyrLeuTyrArgSerTrrpSerGly 120
Db 411 AAGAGGAGCAGTGGCAGTGGATGATGGGGCCATGTATCTGTACAGATCTCTGGTCTGGC 470
Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPhleLeuThr 140
Db 471 AAGTCCATGGTGGGAAACAAGCATGTGCTGAGATGAGCTCCATTAACAACACTTTTAACT 530
Qy 141 TrpSerSerGlnGluCysAsnLysArgGlnHisPheLeuGlyLysTyrArgPro 158
Db 531 TGGAGCAGCAAGAAATGCACAGCCACACACTTCTCTGTGCAAGTACCGACCA 584

RESULT 5
US-09-146-969-1
; Sequence 1, Application US/09146969
; Patent No. 6228585
; GENERAL INFORMATION:
; APPLICANT: Dieckgraebe, Brian K.
; TITLE OF INVENTION: Gene Markers for Chronic Mucosal Injury
; FILE REFERENCE: 04255.75314
; CURRENT APPLICATION NUMBER: US/09/146.969
; CURRENT FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-969-1

Alignment Scores:
Pred. No.: 1.64e-18 Length: 777
Score: 224.00 Matches: 45
Percent Similarity: 52.59% Conservative: 26
Best Local Similarity: 33.33% Mismatches: 58
Query Match: 25.51% Indels: 6
DB: 3 Gaps: 4

US-09-525-041-2 (1-158) x US-09-146-969-1 (1-777)

Qy 27 ArgProSerCysAlaProGlyTrrpPheTyrHisLysSerAsnCysTyrGlyTyrPheArg 46
Db 179 CGGATCAGCTGCCAGAGCCAGCAATGCCTCTCTACTGCTACTACTTAAATGAA 238
Qy 47 LysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSerTyrGlyAsnGlyAlaHis 66
Db 239 GACCTGAGACCTGGTGTATGAGATCTCTATTGCCAGAACATGAATTCGGGC---AAC 295
Qy 67 LeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAlaGluTyrIle-----Ser 84
Db 296 CTGGTGTCTGTGCTCACCAGCGAGGTCCTTTGTGGCTCTCACTGATTAAGGAGAT 355
Qy 85 GlyTyrGlnArgSerGlnProIleTrrpIleGlyLeuHisAspProGlnLysArgGlnGln 104
Db 356 GGCACTGATGACTTCAAT--GTCTGATGGCTCTCCATGACCCCAAAAAGAACCGCGC 412
Qy 105 TrpGlnTrrpIleAspGlyAlaMetTyrLeuTyrArgSerTrrp-----SerGlyLysSer 122
Db 413 TGGCACTGGAGCAGTGGTCT 472
Qy 123 MetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPhleLeuThrTrrpSer 142
Db 473 AGTGTAACTCTGCTACTGTGTGAGCTGACCTCAAGCAGACAGGATTCAGAAATGGAAG 532
Qy 143 SerAsnGluCysAsnLysArgGlnHisPheLeuGlyLysTyrArg 157
Db 533 GATGTGCTCTGTGAAGACAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 577

RESULT 6

US-08-464-637-1
; Sequence 1, Application US/08464637
; Patent No. 5834214
; GENERAL INFORMATION:
; APPLICANT: Iovanna, Juan-Lucio
; APPLICANT: Dagorn, Jean-Charles
; APPLICANT: Keim, Volker
; APPLICANT: Sarles, Jacques
; TITLE OF INVENTION: Detection of Pancreatitis-Associated
; TITLE OF INVENTION: Protein for diagnosis of Cystic Fibrosis or Pancreatic
; TITLE OF INVENTION: Disease (as amended).
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,637
; FILING DATE: 30-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 2121-107P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 797 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 43..567 /product= "human"
; OTHER INFORMATION: pancreatitis-associated protein"
; OTHER INFORMATION: /note= "see, Fig. 3"

US-08-464-637-1

Alignment Scores:
Pred. No.: 4.01e-18 Length: 797
Score: 221.00 Matches: 53
Percent Similarity: 50.62% Conservative: 29
Best Local Similarity: 32.72% Mismatches: 64
Query Match: 25.17% Indels: 16
DB: 2 Gaps: 7

US-09-525-041-2 (1-158) x US-08-464-637-1 (1-797)

Qy 10 LeuLeuLeuSerCysLeu---AlaLysThrGlyValLeuGlyAspIleIleMetArg--- 27
Db 79 ATGCTGCTTTCCTGCTCATGCTGTCTCAGGTTCAAGGTGAAGAACCCACAGGGAA 138
Qy 28 ---ProSer-----CysAlaProGlyTrrpPheTyrHisLysSerAsnCysTyr 42
Db 139 CTGCCCTCTGCAGGATCGCTGTCCCAAAGGCTCAAGGCTCATGGCTCCCACTGCTAT 198


```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/822,261
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0251 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 797 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 262368
US-09-226-852-5

```

```

Alignment Scores:
Pred. No.: 4.01e-18 Length: 797
Score: 221.00 Matches: 53
Percent Similarity: 50.62% Conservative: 29
Best Local Similarity: 32.72% Mismatches: 64
Query Match: 25.17% Indels: 16
DB: 4 Gaps: 7

```

US-09-525-041-2 (1-158) x US-09-226-852-5 (1-797)

```

Qy 10 LeuLeuSerCysLeu---AlaLysThrGlyValLeuGlyAspIleLeuMetArg--- 27
Db 79 ATGCTGCTTCTCGCTCAGTCTCTGCTCAGGTTCAAGGTGAAGAACCCAGAGGAA 138
Qy 28 ---ProSer-----CysAlaProGlyTrpPheTyrHisLysSerAsnCysTyr 42
Db 139 CTGCGCTCTGCAGGATCCGCTGTCCTCAAGAGGCTCAAGGCTATGCTCCACTGCTAT 198
Qy 43 GlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSerTyrGly 62
Db 199 GCCTTGTTTGTTCACCAAAATCTGCACAGATGCAGATCTGGCTGCCAGAGCGGCC 258
Qy 63 AsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAlaGluTyr 82
Db 259 TCTGGA---AACCTGGTGTCTGCTCAGTGGGCTCAGGATCCTTCGTGCTCCTCGT 315
Qy 83 IleSerGlyTyrGlnArgSerGlnPro---IleTrpIleGlyLeuHisAspProGlnLys 101
Db 316 GTGAAGAGCATTGTGTAACAGCTACTCATACGCTCTGGATTGGCTCCATGACCCACACAG 375
Qy 102 ArgGlnGln-----TrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArg 116
Db 376 GGCACCGAGCCCAATGGAGAGGTTGGAGTGGAGTAGCAGATGATGTGATGAATATCTTT 435
Qy 117 SerTrp-----SerGlyLysSerMetGlyLysAsnLysHisCysAlaGluMetSerSer 134
Db 436 GCATGGGAGAGAAATCCCTCCACCATCTCAAGCCCGCCCACTGTGGAGCCTGTCGAGA 495
Qy 135 AsnAsnAsnPheLeuThrTrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCys 154
Db 496 AGCAGCAGATTCTTGAGTGGAGAGATTATTAACGTGTAATGTGAGGTTACCCCTATGCTGC 555
Qy 155 LysTyr 156
Db 556 AAGTTC 561

```

RESULT 9

```

US-07-778-156-4
Sequence 4, Application US/07778156
Patent No. 5436169
GENERAL INFORMATION:

```

```

APPLICANT: IOVANNA, JUAN-LUCIO
APPLICANT: KEIM, VOLKER
APPLICANT: DAGORN, JEAN-CHARLES
TITLE OF INVENTION: PROTEIN ASSOCIATED WITH ACUTE
TITLE OF INVENTION: PANCREATITIS, AGENTS FOR THE DIAGNOSIS OF ACUTE
TITLE OF INVENTION: PANCREATITIS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/778,156
FILING DATE: 19911219
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5436169man P.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 354-012-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 522 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: pancreas
US-07-778-156-4

```

```

Alignment Scores:
Pred. No.: 5.02e-18 Length: 522
Score: 218.00 Matches: 53
Percent Similarity: 50.31% Conservative: 28
Best Local Similarity: 32.92% Mismatches: 64
Query Match: 24.83% Indels: 16
DB: 1 Gaps: 7

```

US-09-525-041-2 (1-158) x US-07-778-156-4 (1-522)

```

Qy 10 LeuLeuSerCysLeu---AlaLysThrGlyValLeuGlyAspIleLeuMetArg--- 27
Db 37 ATGCTGCTTCTCGCTCAGTCTCTGCTCAGGTTCAAGGTGAAGAACCCAGAGGAA 96
Qy 28 ---ProSer-----CysAlaProGlyTrpPheTyrHisLysSerAsnCysTyr 42
Db 97 CTGCGCTCTGCAGGATCCGCTGTCCTCAAGAGGCTCAAGGCTATGCTCCACTGCTAT 156
Qy 43 GlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSerTyrGly 62
Db 157 GCCTTGTTTGTTCACCAAAATCTGCACAGATGCAGATCTGGCTGCCAGAGCGGCC 216
Qy 63 AsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAlaGluTyr 82
Db 217 TCTGGA---AACCTGGTGTCTGCTCAGTGGGCTCAGGATCCTTCGTGCTCCTCGT 273
Qy 83 IleSerGlyTyrGlnArgSerGlnPro---IleTrpIleGlyLeuHisAspProGlnLys 101
Db 274 GTGAAGAGCATTGTGTAACAGCTACTCATACGCTCTGGATTGGCTCCATGACCCACACAG 333

```

QY 43 GlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSerTyrGly 62
Db 199 GCCTTGTCTTTTGTACCAAAATCCTGCAGATGCAGATCTGCCTGCCAGAGGGGCC 258
QY 63 AsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAlaGluTyr 82
Db 259 TCTGGA--AACTGGTGTCTGTCTAGTGGGCTGAGGATCCTTCTGTCCTCCCTG 315
QY 83 IleSerGlyTyrGlnArgSerGlnPro---IleTrpIleGlyLeuHisAspProGlnLys 101
Db 316 GTGAGAGAGCATGTTGTAACAGCTACTCATAGCTCGATTGGCTCCATGACCCACACAG 375
QY 102 ArgGlnGln-----TrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArg 116
Db 376 GGCACCGAGCCCAATCGTCCACCAATCTCAAGCCCGGCCACTGTGCGAGCCTGTGAGA 495
QY 117 SerTrp-----SerGlyLysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSer 134
Db 436 GCATGGAGAGAAATCCCTCCACCACTCAAGCCCGGCCACTGTGCGAGCCTGTGAGA 495
QY 135 AsnAsnAsnPheLeuThrTrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCys 154
Db 496 AGCACAGCATTTCTGAGGTGGAAAGATTATACTGTATGTGAGGTACCCCTATGCTGTC 555
QY 155 LysTyr 156
Db 556 AAGTTC 561

RESULT 7

US-08-822-261-5
; Sequence 5, Application US/08822261
; Patent No. 5935813
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PANCEATITIS-ASSOCIATED
; TITLE OF INVENTION: NOVEL HUMAN PANCEATITIS-ASSOCIATED
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,261
; FILING DATE: Filed Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0251 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 797 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 262368

US-08-822-261-5

Alignment Scores:

Pred. No.: 4,01e-18 Length: 797
Score: 221.00 Matches: 53
Percent Similarity: 50.62% Conservative: 29
Best local Similarity: 32.72% Mismatches: 64
Query Match: 25.17% Indels: 16
DB: 2 Gaps: 7

US-09-525-041-2 (1-158) x US-08-822-261-5 (1-797)

QY 10 LeuLeuLeuSerCysLeu---AlaLysThrGlyValLeuGlyAspIleIleMetArg--- 27
Db 79 ATGCTGCTTCTCTGCTCATGCTGCTCAGGTTCAAGTTGAAGAAAGAACCCAGAGGAA 138
QY 28 ---ProSer-----CysAlaProGlyTrpPheTyrHisLysSerAsnCysTyr 42
Db 139 CTGCCTCTGCACGATCCGCTGTCCCAAGGCTCCAAAGGCTATGGCTCCCACTGCTAT 198
QY 43 GlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSerTyrGly 62
Db 199 GCCTTGTCTTGTCCACCAATCTCGACAGATGAGATCTGGCTGCAGAGCGGCC 258
QY 63 AsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAlaGluTyr 82
Db 259 TCTGGA--AACTGGTGTCTGTCTAGTGGGCTGAGGATCCTTCTGTCCTCCCTG 315
QY 83 IleSerGlyTyrGlnArgSerGlnPro---IleTrpIleGlyLeuHisAspProGlnLys 101
Db 316 GTGAGAGAGCATTGTGTAACAGCTACTCATAGCTCTGGATTGGCTCCATGACCCACACAG 375
QY 102 ArgGlnGln-----TrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArg 116
Db 376 GGCACCGAGCCCAATGGAGAGGTGGAGTGGAGTAGCAGTAGCATGTGATGAATCTTT 435
QY 117 SerTrp-----SerGlyLysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSer 134
Db 436 GCATGGAGAGAAATCCCTCCACCACTCAAGCCCGGCCACTGTGCGAGCCTGTGAGA 495
QY 135 AsnAsnAsnPheLeuThrTrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCys 154
Db 496 AGCACAGCATTTCTGAGGTGGAAAGATTATACTGTATGTGAGGTACCCCTATGCTGTC 555
QY 155 LysTyr 156
Db 556 AAGTTC 561

RESULT 8

US-09-525-852-5
; Sequence 5, Application US/09226852
; Patent No. 6492499
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PANCEATITIS-ASSOCIATED
; TITLE OF INVENTION: NOVEL HUMAN PANCEATITIS-ASSOCIATED
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/226,852
; FILING DATE:

QY 102 ArgGlnGln-----TrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArg 116
Db 334 GGCACCGAGCCCAATGAGAGGTTGGGAGTGGAGTAGCAGTGTGATGAATACTTT 393
QY 117 SerTrp-----SerGlyLysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSer 134
Db 394 GCATGGGAGAGAAATCCCTCCACCATCTCAAGCCCGGCGCACTGTGCGAGCCTGTGCGAGA 453
QY 135 AsnAsnAsnPheLeuThrTrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCys 154
Db 454 AGCACGACATTCTGAGGGTGGAGAAATATACTGTATGTGAGGTACCCCTATGTCTGC 513
QY 155 Lys 155
Db 514 AAA 516

RESULT 10

US-08-422-166-4
; Sequence 4, Application US/08422166
; Patent No. 5959086
; GENERAL INFORMATION:
; APPLICANT: IOVANNA, JUAN-LUCIO
; APPLICANT: KEIM, VOLKER
; APPLICANT: DAGORN, JEAN-CHARLES
; TITLE OF INVENTION: PROTEIN ASSOCIATED WITH ACUTE
; TITLE OF INVENTION: PANCREATITIS, AGENTS FOR THE DIAGNOSIS OF ACUTE
; TITLE OF INVENTION: PANCREATITIS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,166
FILING DATE: 14-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/778,156
FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5959086man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 354-012-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 522 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: pancreas
US-08-422-166-4

Alignment Scores:

Pred. No.: 5,028-18 Length: 522
Score: 218.00 Matches: 53
Percent Similarity: 50.31% Conservative: 28

Best Local Similarity: 32.92% Mismatches: 64
Query Match: 24.83% Indels: 16
DB: 2 Gaps: 7

US-09-525-041-2 (1-158) x US-08-422-166-4 (1-522)

QY 10 LeuLeuLeuSerCysLeu---AlaLysThrGlyValLeuGlyAspIleLeuMetArg---- 27
Db 37 ATGCTGCTTTCCTGCTCATGCTGCTCAAGTTCAAGGTGAAGAACCCAGAGGGAA 96
QY 28 ---ProSer-----CysAlaProGlyTTPheTyrHisLysSerLysCysTyr 42
Db 97 CTGCCCCCTGTCAGCGAATCGCTGTCCAAAGGCTCCAAGGCTTATGGCTCCCACTGCTAT 156
QY 43 GlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSerTyrGly 62
Db 157 GCCTTGTGTTTGTCCACCAAAATCTGGACAGATGAGATCTGGCTGCCAAGAGCGGCC 216
QY 63 AsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAlaGluTyr 82
Db 217 TCTGGA---AACCTGGTGTCTGCTCAGTGGGCTGAGGGATCCTTCGTCTCCTCTG 273
QY 83 IleSerGlyTyrGlnArgSerGlnPro---IleTTrpIleGlyLeuHisAspProGlnLys 101
Db 274 GTGAAGAGCATGGTAAACAGTACTCATACGTCTGGATTGGGCTCCATGACCCACACAG 333
QY 102 ArgGlnGln-----TrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArg 116
Db 334 GGCACCGAGCCCAATGAGAGGTTGGGAGTGGAGTAGCAGTGTGATGAATACTTT 393
QY 117 SerTrp-----SerGlyLysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSer 134
Db 394 GCATGGGAGAGAAATCCCTCCACCATCTCAAGCCCGGCGCACTGTGCGAGCCTGTGCGAGA 453
QY 135 AsnAsnAsnPheLeuThrTrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCys 154
Db 454 AGCACGACATTCTGAGGGTGGAGAAATATACTGTATGTGAGGTACCCCTATGTCTGC 513
QY 155 Lys 155
Db 514 AAA 516

RESULT 11

US-07-778-156-12
; Sequence 12, Application US/07778156
; Patent No. 5436169
; GENERAL INFORMATION:
; APPLICANT: IOVANNA, JUAN-LUCIO
; APPLICANT: KEIM, VOLKER
; APPLICANT: DAGORN, JEAN-CHARLES
; TITLE OF INVENTION: PROTEIN ASSOCIATED WITH ACUTE
; TITLE OF INVENTION: PANCREATITIS, AGENTS FOR THE DIAGNOSIS OF ACUTE
; TITLE OF INVENTION: PANCREATITIS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/778,156
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5436169man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 354-012-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 522 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: pancreas
; US-07-778-156-12

5/4 3 6/69

Db 436 GCATGGGAGAGAAATCCCTCCACCATCTCAAGCCCGGCACCTGTGCGAGCTGTCGAGA 495
QY 135 AsnAsnAsnPhelLeuThrTrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCys 154
Db 496 AGCAGAGCATCTTCGAGGTGGAAGATTATTAAGTGTGAGGTACCCCTATGTCTGC 555
QY 155 Lys 155
Db 556 AAA 558

RESULT 13
US-08-422-166-12
; Sequence 12, Application US/08422166
; Patent No. 5959086
; GENERAL INFORMATION:
; APPLICANT: IOVANNA, JUAN-LUCIO
; APPLICANT: KEIM, VOLKER
; APPLICANT: DAGORN, JEAN-CHARLES
; TITLE OF INVENTION: PROTEIN ASSOCIATED WITH ACUTE
; TITLE OF INVENTION: PANCREATITIS, AGENTS FOR THE DIAGNOSIS OF ACUTE
; TITLE OF INVENTION: PANCREATITIS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,166
; FILING DATE: 14-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/778,156
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5959086man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 354-012-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 798 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: pancreas
US-08-422-166-12

Alignment Scores:
Pred. No.: 9,48e-18 Length: 798
Score: 218.00 Matches: 53
Percent Similarity: 50.31% Conservative: 28
Best Local Similarity: 32.92% Mismatches: 64
Query Match: 24.83% Indels: 16
DB: 2 Gaps: 7
US-09-525-041-2 (1-158) x US-08-422-166-12 (1-798)

QY 10 LeuLeuLeuSerCysLeu---AlaLysThrGlyValLeuGlyAspIleIleMetArg--- 27
Db 79 ATGCTGCTTCTGCTGCTCATGCTGCTCAAGGTGAAGAGAGAGAGAGAGAGAGAGAGAA 138
QY 28 ---ProSer-----CysAlaProGlyTrpPheTyrHisLysSerAsnCysTyr 42
Db 139 CTGCCCTCTGCACGGATCCGCTGCTCCAAAGGCTCAAAGGCTATGGCTCCCACTGCTAT 198
QY 43 GlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSerTyrGly 62
Db 199 GCCTGTGTTTGTCCACCAAAATCTGGACAGATGCAGATCGCCCTGCCAAGAGCGGCC 258
QY 63 AsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAlaGluTyr 82
Db 259 TCTGGA---AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 315
QY 83 IleSerGlyTyrGlnArgSerGlnPro---IleTrpIleGlyLeuHisAspProGlnLys 101
Db 316 GTGAAGAGCATTTGTTAAACAGCTACTCATAGCTCTGGATGGGCTCCATGACCCACACAG 375
QY 102 ArgGlnGln-----TrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArg 116
Db 376 GCACCGAGCCCAATGGAGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 435
QY 117 SerTrp-----SerGlyLysSerMetGlyGlyAsnLysHisCysAlaGluMetSer 134
Db 436 GCATGGGAGAGAAATCCCTCCACCATCTCAAGCCCGGCACCTGTGCGAGCTGTCGAGA 495
QY 135 AsnAsnAsnPhelLeuThrTrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCys 154
Db 496 AGCAGAGCATTTCTGAGGTGGAAGATTATTAAGTGTGAGGTACCCCTATGTCTGC 555
QY 155 Lys 155
Db 556 AAA 558

RESULT 14
US-09-146-969-2
; Sequence 2, Application US/09146969
; Patent No. 6228585
; GENERAL INFORMATION:
; APPLICANT: Dieckgraebe, Brian K.
; TITLE OF INVENTION: Gene Markers for Chronic Mucosal Injury
; FILE REFERENCE: 04255.75314
; CURRENT APPLICATION NUMBER: US/09/146,969
; CURRENT FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 798
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-969-2

Alignment Scores:
Pred. No.: 9,48e-18 Length: 798
Score: 218.00 Matches: 53
Percent Similarity: 50.31% Conservative: 28
Best Local Similarity: 32.92% Mismatches: 64
Query Match: 24.83% Indels: 16
DB: 3 Gaps: 7

US-09-525-041-2 (1-158) x US-09-146-969-2 (1-798)

QY 10 LeuLeuLeuSerCysLeu---AlaLysThrGlyValLeuGlyAspIleIleMetArg--- 27
Db 79 ATGCTGCTTCTGCTGCTCATGCTGCTCAAGGTGAAGAGAGAGAGAGAGAGAGAGAA 138
QY 28 ---ProSer-----CysAlaProGlyTrpPheTyrHisLysSerAsnCysTyr 42
Db 139 CTGCCCTCTGCACGGATCCGCTGCTCCAAAGGCTCAAAGGCTATGGCTCCCACTGCTAT 198
QY 43 GlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSerTyrGly 62

```
Db 199 GCCTTGTCTTTTGTACCAAAATCTGGACAGATGCAGATCTGGCTGCCAGAGGGGCC 258
Qy 63 AsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAlaGluTyr 82
Db 259 TCTGA---AACCTGGTCTCTGTCTAGTGGGCTGAGGATCTTCTGTCTCTCCCTG 315
Qy 83 IleSerGlyTyrGlnArgSerGlnPro---IleTrpIleGlyLeuHisAspProGlnLys 101
Db 316 GTGAAGACATTTGTTAAGCTACTCATACGCTCTGGATTTGGCTCCATGACCCACACAG 375
Qy 102 ArgGlnGln-----TrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArg 116
Db 376 GGCACCGAGCCCAATGGAGAGGTTGGAGTGGAGTAGCAGTGTATGATGAATTAATTT 435
Qy 117 SerTrp-----SerGlyLysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSer 134
Db 436 GCATGGGAGAGAAATCCCTCACCATCTCAAGCCCGGCCACTGTGGCGAGCTGTGCGAGA 495
Qy 135 AsnAsnAsnPheLeuThrTrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCys 154
Db 496 AGCACAGCATTTCTGAGGTGGAAGATTATACTGTAATGTGAGTTACCCCTATGTCTGC 555
Qy 155 Lys 155
Db 556 AAA 558
```

RESULT 15

```
US-09-525-041-2
; Sequence 6, Application US/09226852
; Patent No. 6492499
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PANCEATITIS-ASSOCIATED
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/226,852
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,261
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0251 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 798 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 189600
US-09-525-041-2
```

```
Alignment Scores:
Pred. No.: 9.48e-18 Length: 798
Score: 218.00 Matches: 53
Percent Similarity: 50.31% Conservative: 28
Best Local Similarity: 32.92% Mismatches: 64
Query Match: 24.83% Indels: 16
DB: 4 Gaps: 7
US-09-525-041-2 (1-158) x US-09-226-852-6 (1-798)
Qy 10 LeuLeuLeuSerCysLeu---AlaLysThrGlyValLeuGlyAspIleLeuMetArg--- 27
Db 79 ATGTGCTTTCTCTGCCTCATGCTGTCTCAGGTTCAAGGTGAAGAACCCACAGAGGAA 138
Qy 28 ---ProSer-----CysAlaProGlyTrpPheTyrHisLysSerAsnCysTyr 42
Db 139 CTGCGCTCTGCACGGATCCGCTGTCCAAAGGCTCCNAGGCCTATGGCTCCACTGCTAT 198
Qy 43 GlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSerTyrGly 62
Db 199 GCCTTGTCTTTGTACCAAAATCTGGACAGATGCAGATCTGGCTGCCAGAGCGGCC 258
Qy 63 AsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAlaGluTyr 82
Db 259 TCTGA---AACCTGGTCTCTGTCTAGTGGGCTGAGGATCTTCTGTCTCTCCCTG 315
Qy 83 IleSerGlyTyrGlnArgSerGlnPro---IleTrpIleGlyLeuHisAspProGlnLys 101
Db 316 GTGAAGACATTTGTTAAGCTACTCATACGCTCTGGATTTGGCTCCATGACCCACACAG 375
Qy 102 ArgGlnGln-----TrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArg 116
Db 376 GGCACCGAGCCCAATGGAGAGGTTGGAGTGGAGTAGCAGTGTATGATGAATTAATTT 435
Qy 117 SerTrp-----SerGlyLysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSer 134
Db 436 GCATGGGAGAGAAATCCCTCACCATCTCAAGCCCGGCCACTGTGGCGAGCTGTGCGAGA 495
Qy 135 AsnAsnAsnPheLeuThrTrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCys 154
Db 496 AGCACAGCATTTCTGAGGTGGAAGATTATACTGTAATGTGAGTTACCCCTATGTCTGC 555
Qy 155 Lys 155
Db 556 AAA 558
```

Search completed: December 31, 2003, 10:41:06
Job time : 79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 31, 2003, 09:11:17 ; Search time 43 Seconds
(without alignments)
353.364 Million cell updates/sec

Title: US-09-525-041-2
Perfect score: 878
Sequence: 1 MASRMRLLLLLSCLAKTGV.....LTWSSNECKRQHFLCKYRP 158
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254.5	29.0	165	2 A47148	reg I, regenerating islet cells - mouse
2	254.5	29.0	165	2 A28351	pancreatic stone protein precursor - rat
3	241.5	27.5	142	2 S78596	ovocleidin - chick
4	240	27.3	166	1 RGHU1B	regenerating islet
5	236.5	26.9	173	2 B47148	reg II, regenerating islet
6	236.5	26.9	174	2 I83377	regenerating prote
7	233	26.5	135	2 A38609	lectin, galactose-
8	231	26.3	174	2 S54979	pancreatitis-assoc
9	231	26.3	175	2 A37194	pancreatic thread
10	228.5	26.0	174	1 A48689	pancreatitis-assoc
11	225.5	25.7	166	1 RGHU1A	regenerating islet
12	224	25.5	166	2 A45751	pancreatic stone p
13	222.5	25.3	172	2 S32489	lectin - Iberian r
14	221	25.2	175	2 A49616	pancreatitis-assoc
15	206.5	23.5	152	2 JC7134	agkisacutacin alph
16	203.5	23.2	131	2 JC5058	bitiscetin alpha c
17	203	23.1	175	2 A41719	pancreatic stone p
18	202.5	23.1	152	2 JC2415	echicetin beta cha
19	202.5	23.1	123	2 JC4690	coagulation factor
20	200.5	22.8	125	2 JC5059	bitiscetin beta ch
21	196.5	22.4	146	2 JC7105	aggregatin beta ch
22	196.5	22.4	146	2 JC4691	coagulation factor
23	194	22.1	175	2 S29822	pancreatitis-assoc
24	193	22.0	330	2 T46256	brevicin - human (
25	193	22.0	912	2 A54423	brevicin precursor
26	191	21.8	883	2 S57653	brevican precursor
27	189	21.5	133	2 A47267	botrocetin alpha c
28	185.5	21.1	125	2 B47267	botrocetin beta ch
29	185.5	21.1	129	2 JC4329	coagulation factor

30	185	21.1	883	2 S49126	brevicin precursor
31	185	21.1	1257	2 S28764	neurocan precursor
32	184.5	21.0	146	2 JC7135	agkisacutacin beta
33	182.5	20.8	301	2 S13165	asialoglycoprotein
34	182	20.7	144	2 PC7027	aggregatin alpha cha
35	181	20.6	291	1 LNHU1	hepatic lectin HI
36	181	20.6	3562	2 A47171	chondroitin sulfat
37	179	20.4	1268	2 S52781	neurocan - mouse
38	179	20.4	2397	1 A55535	versican precursor
39	179	20.4	2409	1 A50979	versican precursor
40	178.5	20.3	311	1 LNHU2A	asialoglycoprotein
41	177.5	20.2	162	1 LNRCL	lectin BRA1-1 prec
42	177.5	20.2	162	1 LNRCL	lectin BRA3-2 prec
43	177	20.2	1643	2 T14274	versican precursor
44	177	20.2	3381	2 T42389	versican precursor
45	172.5	19.6	123	2 B42972	coagulation factor

ALIGNMENTS

RESULT 1

A47148
reg I, regenerating islet cells - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 20-Jun-2000
C:Accession: A47148
R:Unno, M.; Yonekura, H.; Nakagawara, K.; Watanabe, T.; Miyashita, H.; Morizumi, S.; Okajima, T.
J. Biol. Chem. 268, 15974-15982, 1993
A:Title: Structure, chromosomal localization, and expression of mouse reg genes, reg I and reg II, regenerating islet cell lectins.
A:Reference number: A47148; MUID:93340209; PMID:8340418
A:Accession: A47148
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-165 <UN>
A:CROSS-references: GB:D14010; NID:g391771; PIDN:BAA03111.1; PID:g391772
C:Genetics:
A:Introns: 21/1; 60/3; 106/3; 144/1
C:Superfamily: tetranectin; C-type lectin homology
F:35-161/Domain: C-type lectin homology <LCH>
F:35-46,63-161,136-153/Disulfide bonds: #status predicted

Query Match	29.0%	Score 254.5;	DB 2;	Length 165;
Best Local Similarity	31.1%;	Pred. No. 6.6e-18;		
Matches	52;	Conservative 35;	Mismatches 51;	Indels 29; Gaps 7;
QY	10	LLLSCLAKTGVLDIIMP	-----SCAPGWFYHKSNVCYFRLRWSDA	54
DB	8	ILLSCLE-----IVLSPSQQAEDLPGARISCPGSGNAYSYCYFTEDRLTWADA	59	
QY	55	ELECOYNGGAHLASILSLKEASTIAEYI--SGYORSQPIWGLHDPKQKQOWIDGAM	112	
DB	60	DLFCQWNSG-YLVSVLQAEGNPFVASLIKESGTTDAN-VWTGLHDPKRNRRWHSSGSL	117	
QY	113	YLKRSW-SGKMGGNK-HCAEMSSNNFLTWSSNECKRQHPFLCKYR	157	
DB	118	FLYKSWATGSPNNSNRGVCVSLTSNTGYKKWKDDNCDQAQSFVCKFK	164	

RESULT 2

A28351
pancreatic stone protein precursor - rat
N:Alternate names: lithostathine
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 11-May-2000
C:Accession: A28351; A39081; FLO147; S34618
R:Terazono, K.; Yamamoto, H.; Takasawa, S.; Shiga, K.; Yonemura, Y.; Tochino, Y.; Okamoto, J.
J. Biol. Chem. 263, 2111-2114, 1988
A:Title: A novel gene activated in regenerating islets.
A:Reference number: A28351
A:Accession: A28351
A:Molecule type: mRNA
A:Residues: 1-165 <TER>

Query Match 27.5%; Score 241.5; DB 2; Length 142;
Best Local Similarity 33.8%; Pred. No. 1.1e-16;
Matches 46; Conservative 23; Mismatches 56; Indels 11; Gaps 3;

QY 30 CAPQWFYHKNCYGYFRKLNRWSDAELECCQSYGNGAHLASILSLKEASTIAEYI----- 83
DB 5 CGFQWPTPGCGLCGFFRELSWSRAESFCRWRGPGSHLAAVRSAAELRLLAELLNARSGG 64
QY 84 --SGYQRSQPIWIGLHPQKRQWQWIDGAMLYLRW--SGKSMGNGKHCAMSSNNFL 139
DB 65 DSGSEGADGRVIGLHRRPAGSRWRWSDGTAPRFASWHTAKARRGGR-CAALRDDEAFT 123
QY 140 TWSNNECNKQHFLECK 155
DB 124 SWAAPCTERNFAVCK 139

RESULT 4
RGHUB
N;Alternate names: reg-lectin 1-beta precursor - human
N;Contains: pancreatic stone protein (PSP)
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence revision 03-Aug-1995 #text_change 16-Jun-2000
C;Accession: S34591; S42729; A44712
R;Bartoli, C.; Gharib, B.; Giorgi, D.; Sansonetti, A.; Dagorn, J.C.; Berge-Lefranc, J.L.
FEBS Lett. 327, 289-293, 1993
A;Title: A gene homologous to the reg gene is expressed in the human pancreas.
A;Reference number: S34591; MUID:93351647; PMID:8348956
A;Accession: S34591
A;Molecule type: DNA
A;Residues: 1-166 <BAR>
A;Cross-references: GB:L08010; NID:G307368; PIDN:AAA18204.1; PID:G487726
A;Note: this gene appears to be expressed in pancreas and liver
R;Morizumi, S.; Watanabe, T.; Unno, M.; Nakagawara, K.; Suzuki, Y.; Miyashita, H.; Yone
Biochim. Biophys. Acta 1217, 199-202, 1994
A;Title: Isolation, structural determination and expression of a novel reg gene, human re
A;Reference number: S42729; MUID:94153997; PMID:8110835
A;Accession: S42729
A;Molecule type: mRNA
A;Residues: 1-166 <MO2>
A;Cross-references: GB:D16816; NID:G474305; PIDN:BAA04091.1; PID:G474306
A;Accession: A44712
A;Molecule type: DNA
A;Residues: 1-166 <MO2>
A;Cross-references: GB:D17291; NID:G474307; PIDN:BAA04124.1; PID:G474308
C;Comment: This protein is found in pancreatic calculi of mammals. The tryptic-like cleav
C;Genetics:
A;Gene: GDB:REG1B; REG1
A;Cross-references: GDB:342079
A;Map position: 2p12-2p12
A;Introns: 22/1; 61/3; 107/3; 145/1
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: glycoprotein; lectin; pancreas; pyroglutamic acid
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-166/Product: regenerating islet lectin beta #status predicted <MAT>
F;34-166/Product: pancreatic stone protein #status predicted <MAT2>
F;36-162/Domain: C-type lectin homology <LCH>
F;23/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;27/Binding site: carbohydrate (Thr) (covalent) #status predicted
F;33-34/Cleavage site: Arg-Ile (trypsin) #status predicted
F;36-47,64-162,137-154/Disulfide bonds: #status predicted

Query Match 27.3%; Score 240; DB 1; Length 166;
Best Local Similarity 34.3%; Pred. No. 1.8e-16;
Matches 46; Conservative 27; Mismatches 57; Indels 4; Gaps 4;

QY 27 RPSCAPGWFYHKNCYGYFRKLNRWSDAELECCQSYGNGAHLASILSLKEASTIAEYISGY 86
DB 33 RISCPEGTNAYRSYCYIFNEDPETWVDADLYCQNNWNSG-NLVSVLTAQBGAFVASLIKES 91
QY 87 QRQSP-IWIGLHDPKRQWQWIDGAMLYLRW--SGKSMGNGK-HCAEMSSNNNFLTWS 143

Db 92 STDDSNVWGLHDPKPKRRHWSGSLVSYKSWDTGSPSSANAGYCASLTSCSGFKKWKD 151

QY 144 NECNKRQHFLECKYR 157

Db 152 ESCEKKFVFCKPK 165

RESULT 5

reg II, regenerating islet cells - mouse

C:Species: Mus musculus (house mouse)

C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 20-Jun-2000

C:Accession: B47148

J:Unno, M.; Yonekura, H.; Nakagawara, K.; Watanabe, T.; Miyashita, H.; Moriizumi, S.; Okamoto, H. Biol. Chem. 268, 15974-15982, 1993

A:Title: Structure, chromosomal localization, and expression of mouse reg genes, reg I and reg II

A:Reference number: A47148; MUID:93340209; PMID:8340418

A:Accession: B47148

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-173 <UNN>

A:Cross-references: GB:D14011; NID:g391773; PIDN:BAA03112.1; PID:g391774

C:Genetics:

A:Introns: 22/1; 68/3; 114/3; 152/1

C:Superfamily: tetranectin; C-type lectin homology

F:43-169/Domain: C-type lectin homology <LCH>

F:43-54,71-169,144-161/Disulfide bonds: #status predicted

Query Match 26.9%; Score 236.5; DB 2; Length 173;

Best Local Similarity 28.7%; Pred. No. 4.2e-16;

Matches 50; Conservative 33; Mismatches 72; Indels 19; Gaps 5;

QY 1 MASRSMRLLLLSCLAKTGVLDII-----MRPSCAPGWYHKNCYGYPRK 47

Db 1 MAQNYYLILFLMLFSLVSQGVAREDFLEAKDLPKAKINGPEGANNYGSYCYLIED 60

QY 48 LRNWDALCQSYNGNHAHLASILSLKEASTIAEYI--SCYQSQPIWIGLHDPKQKQW 105

Db 61 RLWTGADLFQNM--NAGHLVLSLSQAESNFVSLVKESGTTASN--VMTGLHDPKSNRRW 118

QY 106 QWIDGAMLYRWS--GKSMGNKHKCAEMSSNNFLTWSNECNKRQHFLECKYR 157

Db 119 HWSSGSLFLFKWATGAPSTANRGYCVSLTNTAYKKKWDENCEAQYSPVCKFR 172

RESULT 6

regenerating protein III (reg III) - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 20-Aug-1999

C:Accession: I60296; I63377

R:Suzuki, Y.; Yonekura, H.; Watanabe, T.; Unno, M.; Moriizumi, S.; Miyashita, H.; Okamoto, H. Gene 144, 315-316, 1994

A:Title: Structure and expression of a novel rat RegIII gene.

A:Reference number: I60296; MUID:94314238; PMID:8039722

A:Accession: I60296

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-174 <RES>

A:Cross-references: GB:D23676; NID:g471157; PIDN:BAA04904.1; PID:g471158

A:Accession: I63377

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-174 <RE2>

A:Cross-references: GB:D26078; NID:g471159; PIDN:BAA05071.1; PID:g471160

C:Genetics:

A:Gene: RegIII

A:Introns: 25/1; 64/3; 110/3; 153/1

C:Superfamily: tetranectin; C-type lectin homology

F:39-170/Domain: C-type lectin homology <LCH>

Query Match 26.9%; Score 236.5; DB 2; Length 174;

Best Local Similarity 35.3%; Pred. No. 4.2e-16;

Matches 49; Conservative 25; Mismatches 56; Indels 9; Gaps 4;

QY 27 RPSCAPGWYHKNCYGYPRKLRNWSDALECCSYNGNHAHLASILSLKEASTIAEYISG- 85

Db 36 RTSCPMGSKAYRSYCYTLVTTLKSWFOADLACQKPSG-HLVSLISGSGEASFVSSLVTCR 94

QY 86 YORSQPIWIGLHDPKQKQ-----QWIDGAMLYRWSG--KSMGNKHKCAEMSSNNNP 138

Db 95 VNNQDIWIGLHDPKQKQNGGWSNSDLNLYLNWDGDPSTVNRNGCGSLTATSEF 154

QY 139 LTWSSNECNKRQHFLECKYR 157

Db 155 LKMGDHHCDVELPFVCKPK 173

RESULT 7

A38609

lectin, galactose-specific - western diamondback rattlesnake

C:Species: Crotalus atrox (western diamondback rattlesnake)

C>Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 11-Aug-1995

C:Accession: A38609

R:Hirabayashi, J.; Kusunoki, T.; Kasai, K. J. Biol. Chem. 266, 2320-2326, 1991

A:Title: Complete primary structure of a galactose-specific lectin from the venom of the rattlesnake

A:Reference number: A38609; MUID:91115849; PMID:1989986

A:Accession: A38609

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-135 <HIR>

C:Superfamily: tetranectin; C-type lectin homology

F:3-131/Domain: C-type lectin homology <LCH>

F:3-14,31-131,106-123/Disulfide bonds: #status predicted

Query Match 26.5%; Score 233; DB 2; Length 135;

Best Local Similarity 34.4%; Pred. No. 7.2e-16;

Matches 45; Conservative 18; Mismatches 64; Indels 4; Gaps 2;

QY 29 SCAPGWYHKNCYGYPRKLRNWSDALECCSYNGNHAHLASILSLKEASTIAEYISGYOR 88

Db 2 NCPLDMLPMLGCLYKIFNLQKTWEDAEFRCYKKGCHLASPHRYGESLEIAEYISDYHK 61

QY 89 SQP-TWIGLHDPKQKQWIDGAMLYRWSGKS---MGNKHKCAEMSSNNNPFLTWSN 144

Db 62 GQENWIGLRKKKDPFSEWTDRTSCTDYLTDKNQPDHYQNKFCVELVSLTGYRLWMDQ 121

QY 145 ECNKRQHFLECK 155

Db 122 VCESKDAFLCQ 132

RESULT 8

pancreatitis-associated protein PAP-3 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 20-Aug-1999

C:Accession: S54979; S43438

R:Dusetti, N.J.; Frigerio, J.M.; Szpirer, C.; Dagorn, J.C.; Iovanna, J.L. Biochem. J. 307, 9-16, 1995

A:Title: Cloning, expression and chromosomal localization of the rat pancreatitis-associated protein PAP-3

A:Reference number: S54979; MUID:95234061; PMID:7717998

A:Accession: S54979

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-174 <DUS>

A:Cross-references: EMBL:U09193; NID:g483931; PIDN:AAA79231.1; PID:g483932

R:Frigerio, J.M.; Dusetti, N.J.; Garrido, P.; Dagorn, J.C.; Iovanna, J.L. Biochim. Biophys. Acta 1216, 329-331, 1993

A:Title: The pancreatitis associated protein III (PAP III), a new member of the PAP gene family

A:Reference number: S43438; MUID:94060113; PMID:8241280

A:Accession: S43438

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-174 <FRI>

A; Cross-references: GB:I20869; GB:S67496; NID:9463279; PIDN:AAA41809.1; PID:9463280
C; Geneticks:
A; Introns: 26/1; 65/3; 111/3; 153/1
C; Superfamily: tetranectin; C-type lectin homology
F; 40-170/Domain: C-type lectin homology <LCH>

Query Match 26.3%; Score 231; DB 2; Length 174;
Best Local Similarity 32.4%; Pred. No. 1.5e-15;
Matches 55; Conservative 26; Mismatches 73; Indels 14; Gaps 5;

QY 1 MASRSWELLLLSLAKTGVLDG-----IIMPSCAPGWFYHKNSCYGFERKLRNWSDA 54
DB 5 VALTTSWMLLSLMLLSVQGDGDAEDVPTSRISCPKSGRAYSGSYCYALFVSVKSWFDA 64
QY 55 ELECQSYNGAHLASILSLKEASTIAEYI-SGYQRSQPIWIGLHDPKQKQ-----WQWI 108
DB 65 DLACQKPSG-HLVSVLSGSEAFVSVLSLKSSGNSGQNVWIGLHDFPLGQEPNRCGWEWS 123
QY 109 DGAMYLYRSW-SGKSMGNKHCIAEMSNNNFLTWSNECKNRQHFLCKYR 157
DB 124 NADVMYFNWETNPSSVSGSHCGTLTRASGFLRWRENNCISLPYVCKFK 173

RESULT 9
A37194
pancreatic thread protein precursor - bovine
C; Species: Bos primigenius taurus (cattle)
C; Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 20-Aug-1999
C; Accession: A37194; A53897
R; de la Monte, S.M.; Ozturk, M.; Wands, J.R.
J. Clin. Invest. 86, 1004-1013, 1990
A; Title: Enhanced expression of an exocrine pancreatic protein in Alzheimer's disease
A; Reference number: A37194; MUID:90368981; PMID:2394826
A; Accession: A37194
A; Molecule type: mRNA
A; Residues: 1-175 <DB>
A; Cross-references: GB:M59794; NID:g163648; PIDN:AAA30750.1; PID:g163649
R; Cai, L.; Harris, W.R.; Marshak, D.R.; Gross, J.; Crabb, J.W.
J. Protein Chem. 9, 623-632, 1990
A; Title: Structural analysis of bovine pancreatic thread protein.
A; Reference number: A53897; MUID:91197388; PMID:2085387
A; Accession: A53897
A; Molecule type: protein
A; Residues: 38-138; 141-175 <CAI>
C; Comment: The purified protein undergoes a reversible globule-fibril transformation and
C; Superfamily: tetranectin; C-type lectin homology
C; Keywords: disulfide bond; extracellular protein; pancreas
F; 38-138/Product: pancreatic thread protein chain A #status experimental <ACH>
F; 40-171/Domain: C-type lectin homology <LCH>
F; 141-175/Product: pancreatic thread protein chain B #status experimental <BCH>

Query Match 26.3%; Score 231; DB 2; Length 175;
Best Local Similarity 30.9%; Pred. No. 1.5e-15;
Matches 51; Conservative 30; Mismatches 62; Indels 22; Gaps 6;

QY 10 LLSLAKTGVLDG-----IIMPSCAPGWFYHKNSCYGFERKLRNWSDALECO 59
DB 13 MLLSCLM---LLSQIQENSQKELPSARIICPSGSMAYRSHCYALFKPTWWDADIACQ 69
QY 60 SYNGAHLASILSLKEASTIAEYISGYQRSQ-PIWIGLHDPKQKQ-----WQWIDGAMY 113
DB 70 KRPFG-HLVSVLSGAESEFVASLVRNNLTQSDIWLGHDPTEGSEANAGGWEWISNDVL 128
QY 114 LYRSW---SGKSMGNKHCIAEMSNNNFLTWSNECKNRQHFLCKY 156
DB 129 NYAVETDPAIISPGYCGSLRSSGYLKRWDHNCNLNLPYVCKF 173

RESULT 10
A48689
pancreatitis-associated protein PAP-2 - rat
C; Species: Rattus norvegicus (Norway rat)
C; Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A48689
R:Frigerio, J.M.; Duseti, N.J.; Keim, V.; Dagorn, J.C.; Iovanna, J.L.
Biochemistry 32, 9236-9241, 1993
A>Title: Identification of a second rat pancreatic-islet-associated protein. Messenger RNA cJ
A:Reference number: A48689; MUID:93378971; PMID:8369291
A:Accession: A48689
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-174 <PRI>
A:Cross-references: GB:I10229; NID:g409014; PIDN:AAA02980.1; PID:g409015
C:Superfamily: tetranectin; C-type lectin homology
E:3/170/Domain: C-type lectin homology <LCH>
F:39-50,67-170,145-162/Disulfide bonds: #starch predicted

Query Match 26.0%; Score 228.5; DB 1; Length 174;
Best Local Similarity 34.5%; Pred. No. 2.6e-15;
Matches 48; Conservative 25; Mismatches 57; Indels 9; Gaps 4;

Qy 27 RPSCAPGWYHKSNICYGFKLRNWSDAELRCQSYNGNAHLASILSLKEASTIAEYISG- 85
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 36 RTSCPMGSKAYRSYCITLVTTLSKFQADLACQKRPSG-HLVSVILSGGEASFVSSLVTGR 94
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Qy 86 YQRSOPTWIGLHDPKQQ-----WQITDGAWLYRWSG--KSMGGNKHCHEMSSNNNF 138
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 95 VNNDIWIWLHDDPTMGQPNGGWEWSNSDLNLVNWGDPPSTVNRNGCSLTATSEF 154
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Qy 139 LTWSSNECNKRQHFLCKYR 157
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 155 LKWGDHCDVELPFVCKFK 173
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

RESULT 11
RGHUIA
regenerating islet lectin 1-alpha precursor [validated] - human
N:Alternate names: lithostathine; pancreatic thread protein (PTP); reg I protein; reg1-aJ
C:Contains: pancreatic stone protein (PSP)
C:Species: Homo sapiens (man)
C>Date: 31-Mar-1990 #sequence revision 03-Aug-1995 #text change 08-Dec-2000
C:Accession: A35197; S29351; S12950; S02767; S02419; S00113; S01471; A25246
R:Watanabe, T.; Yonekura, H.; Terazono, K.; Yamamoto, H.; Okamoto, H.
J. Biol. Chem. 265, 7432-7439, 1990
A>Title: Complete nucleotide sequence of human reg gene and its expression in normal and product of the gene.
A:Reference number: A35197; MUID:90237042; PMID:2332435
A:Accession: A35197
A:Molecule type: DNA
A:Residues: 1-166 <WAT>
A:Cross-references: GB:J05412
R:Terazono, K.; Yamamoto, H.; Takasawa, S.; Shiga, K.; Yonemura, Y.; Tochino, Y.; Okamoto, J.
J. Biol. Chem. 263, 2111-2114, 1988
A>Title: A novel gene activated in regenerating islets.
A:Reference number: A92704; MUID:88115343; PMID:2963000
A:Accession: B28351
A:Molecule type: mRNA
A:Residues: 1-166 <TER>
A:Cross-references: GB:M18963; NID:g190978; PIDN:AAA36558.1; PID:g190979
R:Itoh, T.; Tsuzuki, H.; Katoh, T.; Teraoka, H.; Matsumoto, K.; Yoshida, N.; Terazono, K.
FEBS Lett. 272, 85-88, 1990
A>Title: Isolation and characterization of human reg protein produced in Saccharomyces cerevisiae.
A:Reference number: S12950; MUID:91032149; PMID:2226837
A:Accession: S12950
A:Molecule type: protein
A:Residues: 23-52,160-166 <ITO>
A>Note: sequence determined from protein isolated after human cdna sequence was cloned ar R:de Caro, A.M.; Adrich, Z.; Fournet, B.; Capon, C.; Bonicel, J.J.; de Caro, J.D.; Rovery Biochim. Biophys. Acta 994, 281-284, 1989
A>Title: N-terminal sequence extension in the glycosylated forms of human pancreatic stor A:Reference number: S02767; MUID:89150292; PMID:2493268
A:Accession: S02767
A:Molecule type: protein
A:Residues: 23-47 <DEC>
R:Ruimi, P.; de Caro, J.; Bonicel, J.; Rovery, M.; de Caro, A.
FEBS Lett. 229, 171-174, 1988

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-166 <GIO>

A:Cross-references: GB:M27190; NID:G623412; PIDN:AAA60546.1; PID:G623413

C:Superfamily: tetranectin; C-type lectin homology

F:36-162/Domain: C-type lectin homology <LCH>

Query Match 25.5%; Score 224; DB 2; Length 166;

Best Local Similarity 33.3%; Pred. No. 7e-15;

Matches 45; Conservative 26; Mismatches 58; Indels 6; Gaps 4

Qy 27 RPSCAPGFVHKSNICYGFKRLNWSDAELECQSYGNGAHLASILSLKEASTIAEYI--S 84

Db 33 RISCPETNAYRSYCYFENSDRETWTVDADLYCQNMNSG-NLVSVLTQAEGAFVASLIKES 91

Qy 85 GYKRSQBIWGLDHPQKROQOWIDGAMLYRSW--SGKSMGGNKHCAEMSSNNNFLTWS 142

Db 92 GTDDFN-VWGLDHPKRNRRHHWSSGSLVYSKMGICAPSVNPGYCVSLTSSITGFQKWK 150

Qy 143 SNECNKRQHFCLKYR 157

Db 151 DVPCEDKFSPVCKPK 165

RESULT 13

S32489

lectin - Iberian ribbed newt

C:Species: Pleurodeles waltlilii (Iberian ribbed newt)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000

C:Accession: S32489; S28530

R:Tiffoche, C.; Chesnel, A.; Jegou, P.; le Pennec, J.P.

Eur. J. Biochem. 213, 901-907, 1993

A:Title: Isolation and characterization of a cDNA clone encoding a Pleurodeles

A:Reference number: S32489; MUID:93279340; PMID:8504829

A:Accession: S32489

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-172 <TIF>

A:Cross-references: EMBL:X69062; NID:G64257; PIDN:CAA48800.1; PID:G64258

C:Superfamily: tetranectin; C-type lectin homology

F:37-164/Domain: C-type lectin homology <LCH>

Query Match 25.3%; Score 222.5; DB 2; Length 172;

Best Local Similarity 34.3%; Pred. No. 1e-14;

Matches 46; Conservative 22; Mismatches 55; Indels 11; Gaps 4

Qy 30 CARGWFVHKSNICYGFKRLNWSDAELECQSYGNGAHLASILSLKEASTIAEYISGVQRS 89

Db 37 CTTPGDWCHFNYSKYIIPNAKSWTDAEPYCQKLYPGAHLASIHSEDENDFLTEITFPKNSN 96

Qy 90 QP-IWGLDHPQKROQOWIDGAMLY--RSWSGKSMGGNKHCAEMSSNNNFLT--WS 142

Db 97 YPVVWGGSDCYKDRSFVWTDGSDWDYQKRWQEPSTGGRPCIDF----NFVTPGLWN 152

Qy 143 SNECNKRQHFCLKY 156

Db 153 DEHCDQKFPFICKY 166

RESULT 14

A49616

pancreatitis-associated protein precursor - human

N:Alternate names: C-type lectin; pancreatic stone protein homolog HIP

C:Species: Homo sapiens (man)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 21-Jul-2000

C:Accession: A49616; A49931; S29821; S48197; I55580

R:Dusetti, N.J.; Frigerio, J.M.; Fox, M.F.; Swallow, D.M.; Dagorn, J.C.; Iovannin

Genomics 19, 108-114, 1994

A:Title: Molecular cloning, genomic organization, and chromosomal localization

A:Reference number: A49616; MUID:94245143; PMID:8188210

A:Accession: A49616

A:Molecule type: DNA

A:Residues: 1-175 <DUS>

GenCore version 5.1.6
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 31, 2003, 09:11:17 ; Search time 68 Seconds
(without alignments)
599.593 Million cell updates/sec

Title: US-09-525-041-2
Perfect score: 878
Sequence: 1 MASRSWRLLLLLSCLAKTGV.....LTWSSNECNKQHFCKYRP 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	878	100.0	158	4	Q9BYZ8
2	600.5	68.4	157	11	Q9D8G5
3	594.5	67.7	157	11	Q9D8S8
4	544	62.0	113	4	Q8NER7
5	307.5	35.0	134	4	Q8NER6
6	235	26.8	126	11	Q8C6F9
7	234	26.7	135	13	Q9PSN0
8	230.5	26.3	154	13	Q8JIV9
9	229	26.1	135	13	Q9PSM4
10	227.5	25.9	146	11	Q9CVF4
11	227	25.9	158	13	Q9OW18
12	227	25.9	158	13	Q9OW17
13	221.5	25.2	155	13	Q8JIV8
14	214	24.4	158	13	Q9OW16
15	213	24.3	125	13	Q9PSM8
16	211.5	24.1	152	13	Q9DG39

17	205.5	23.4	146	13	Q9DG31	Q9dg31 agkistrodon
18	205	23.3	163	13	Q9AXR8	Q9axr8 anguilla ja
19	204.5	23.3	152	13	Q9DEF9	Q9def9 agkistrodon
20	204.5	23.3	152	13	Q9IAM1	Q9iam1 agkistrodon
21	204.5	23.3	152	13	Q8JIW0	Q8jiw0 agkistrodon
22	203.5	23.2	154	13	Q9YGN5	Q9ygn5 agkistrodon
23	203	23.1	155	13	Q9DEA2	Q9dea2 agkistrodon
24	202.5	23.1	146	13	Q9U0J0	Q9u0j0 trimeresu
25	200.5	22.8	158	13	Q8JGT7	Q8jgt7 trimeresu
26	199	22.7	155	13	Q8JIV6	Q8jiv6 agkistrodon
27	197.5	22.5	158	13	Q8AVS7	Q8avs7 trimeresu
28	196.5	22.4	131	13	Q9PSM9	Q9psm9 echis carin
29	196.5	22.4	146	13	Q9I840	Q9i840 agkistrodon
30	194.5	22.2	146	13	Q8JIV7	Q8jiv7 agkistrodon
31	193.5	22.0	124	13	Q98SM5	Q98sm5 agkistrodon
32	193	22.0	330	4	Q9NT67	Q9nt67 homo sapien
33	193	22.0	911	4	Q9HBK4	Q9hbk4 homo sapien
34	193	22.0	911	4	Q96FP7	Q96fp7 homo sapien
35	193	22.0	911	4	Q96GW7	Q96gw7 homo sapien
36	193	22.0	911	4	Q8TBB9	Q8tbb9 homo sapien
37	192.5	21.9	146	13	Q8AYA4	Q8aya4 agkistrodon
38	191.5	21.8	146	13	Q9DEF8	Q9def8 agkistrodon
39	190	21.6	146	13	Q8JGT6	Q8jgt6 trimeresu
40	189	21.5	292	4	Q14538	Q14538 homo sapien
41	189	21.5	316	4	Q8IUN9	Q8iun9 homo sapien
42	186	21.2	148	13	Q8AVS8	Q8avs8 trimeresu
43	185.5	21.1	146	13	Q9IAM0	Q9iam0 agkistrodon
44	185.5	21.1	158	13	Q8UVC6	Q8uvc6 agkistrodon
45	185.5	21.1	158	13	Q8AYA5	Q8aya5 agkistrodon

ALIGNMENTS

RESULT 1

Q9BYZ8	PRELIMINARY;	PRT;	158 AA.
AC	Q9BYZ8		
DT	01-JUN-2001 (Tremblrel. 17, Created)		
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)		
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)		
DE	Regenerating gene type IV (Gastrointestinal secretory protein GISP)		
DE	(REG-like protein).		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]_TaxID=9606;		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21210973; PubMed=11311942;		
RA	Hartup J.C., Zhang H., Bonaudo M.F., Soares M.B., Dieckgraebe B.K.;		
RT	"Isolation and characterization of a cDNA encoding a novel member of		
RT	the human regenerating protein family; Reg IV(1).";		
RL	Biochim. Biophys. Acta 1518:287-293(2001).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Colon cancer;		
RA	Violette S., Fesfor E., Dussaux E., Citadelle D., Chambaz J.,		
RA	Lacasa M., Lesuffleur T.;		
RT	"The new intestinal Reg IV gene is involved in drug-resistance of		
RT	colon tumor cells.";		
RL	.Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Colon;		
RA	Strausberg R.;		
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RA	Lin W.-C.;		
RT	"Identification of gastrointestinal secretory protein (GISP), a new		
RT	member of lithostathine gene family.";		
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.		

```

RN [5]
RP SEQUENCE FROM N.A.
RA Kammarainen M., Heiskala K., Heiskala M., Andersson L.C.;
RT "RELp, a novel human REG-like protein.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY007243; AAG02562.1; -
DR EMBL; AF345934; AAX59869.1; -
DR EMBL; BC017089; AAH17089.1; -
DR EMBL; AF254415; AAK48435.1; -
DR EMBL; AY126670; AAM95598.1; -
DR HSSP; P05451; 1QDD.
DR InterPro; IPR001304; Lactin C.
DR InterPro; IPR003990; Pancreatins_ac.
DR Pfam; PF00059; lectin C; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR PROSITE; PS00615; C_TYPE_LACTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LACTIN_2; 1.
KW signal.
FT SIGNAL 1 22 POTENTIAL.
SQ SEQUENCE 158 AA; 18230 MW; 7308849CBBDE593E CRC64;

Query Match 100.0%; Score 878; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.8e-87;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASSMRLLLLSCLAKTGLVDIIMRPSCAPGWFYHKSNICYGYFRKLRNWSDALEECQ 60
DB 1 MASSMRLLLLSCLAKTGLVDIIMRPSCAPGWFYHKSNICYGYFRKLRNWSDALEECQ 60
QY 61 YNGAHLASILSLKEASTIAEYISGYQSORPIWIGLHDPKQKQOWIDGAMVLYRSWSG 120
DB 61 YNGAHLASILSLKEASTIAEYISGYQSORPIWIGLHDPKQKQOWIDGAMVLYRSWSG 120
QY 121 KSMGKHKCAEMSSNNFLTWSSNECNKROHFLCKYRP 158
DB 121 KSMGKHKCAEMSSNNFLTWSSNECNKROHFLCKYRP 158

```

```

RT "Functional annotation of a full-length mouse cdna collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE=Colon;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK008049; BAB25429.1; -
DR EMBL; BC019465; AAH19465.1; -
DR HSSP; P05451; 1QDD.
DR MGD; MGI:1914959; 2010002L15RIK.
DR InterPro; IPR001304; Lactin C.
DR InterPro; IPR003990; Pancreatins_ac.
DR Pfam; PF00059; lectin C; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LACTIN_2; 1.
SQ SEQUENCE 157 AA; 18398 MW; F3981722BBD83968 CRC64;

Query Match 68.4%; Score 600.5; DB 11; Length 157;
Best Local Similarity 66.2%; Pred. No. 4.4e-57;
Matches 104; Conservative 25; Mismatches 27; Indels 1; Gaps 1;

QY 1 MASSMRLLLLSCLAKTGLVDIIMRPSCAPGWFYHKSNICYGYFRKLRNWSDALEECQ 60
DB 1 MASKGVRLLLLSWAGPEVLS-D-ILRPSCAPGWFYRSHCYGYFRKLRNWSHALEECQ 59
QY 61 YNGAHLASILSLKEASTIAEYISGYQSORPIWIGLHDPKQKQOWIDGAMVLYRSWSG 120
DB 60 YNGSHSLASVLNKEASVISKITYQRLNPLVWIGLHDPKQKQOWIDGSTNLRYRNP 119
QY 121 KSMGKHKCAEMSSNNFLTWSSNECNKROHFLCKYR 157
DB 120 RTKSEARHCAEMNPKDKFLTNKNGCANRQHFLCKYK 156

RESULT 3
Q9D858 PRELIMINARY; PRT; 157 AA.
ID Q9D858; AC Q9D858;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 2010002L15RIK protein.
GN 2010002L15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Small intestine;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Azakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cdna collection.";
RL Nature 409:685-690(2001).

```

```
DR EMBL; AK008438; BAB25669.1; -.
DR HSP; P05451; 10DD.
DR MGD; MG1:1914959; 2010002L15R1K.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR003990; Pancreatiss_ac.
DR Pfam; PF00059; lectin C; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE LECTIN 2; 1.
SQ SEQUENCE 157 AA; 18474 MW; FD96F36CFB989368 CRC64;

Query Match 67.7%; Score 594.5; DB 11; Length 157;
Best Local Similarity 65.6%; Pred. No. 1.9e-56;
Matches 103; Conservative 25; Mismatches 28; Indels 1; Gaps 1;

QY 1 MASRSRLRLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFYFKLRNWSDAELECOS 60
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MAYKGVRLJLLLSWAGPEVLSD-IIRPSCAPGWFYHKSNCYGFYFKLRNWSHAELECOS 59
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 61 YNGAHLASILSLKEASTIAEYISGYQSORQPIWIGLHDPQKQOWIDGAMLYLRSWSG 120
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 60 YNGSHLASVNLQKEASVISKYITQYRNLPVWIGLHDPQKQOWIDGSTNLYRRWNP 119
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 121 KSGGNGKHAEMSSNNFLTWSSNECKRQHFLCKYR 157
   ::|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 120 RTKSEARHCAEMPNKDKFLTWKNGCANRQHFLCKYK 156
   ::|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 4
Q8NER7 PRELIMINARY; PRT; 113 AA.
AC Q8NER7
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE REG-like protein splice variant 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kamarainen M., Heiskala K., Heiskala M., Andersson L.C.;
RT "REL", a novel human REG-like protein."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV126671; AAM95599.1; -.
DR InterPro; IPR001304; Lectin C.
DR Pfam; PF00059; lectin C; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR PROSITE; PS50041; C_TYPE LECTIN 2; 1.
SQ SEQUENCE 113 AA; 12832 MW; A2E9DF1A729C78DA CRC64;

Query Match 62.0%; Score 544; DB 4; Length 113;
Best Local Similarity 99.0%; Pred. No. 3.8e-51;
Matches 101; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSRLRLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFYFKLRNWSDAELECOS 60
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MASRSRLRLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFYFKLRNWSDAELECOS 60
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 61 YNGAHLASILSLKEASTIAEYISGYQSORQPIWIGLHDPQK 102
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 YNGAHLASILSLKEASTIAEYISGYQSORQPIWIGLHDPQK 102
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 5
Q8NER6 PRELIMINARY; PRT; 134 AA.
AC Q8NER6
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE REG-like protein splice variant 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kamarainen M., Heiskala K., Heiskala M., Andersson L.C.;
RT "REL", a novel human REG-like protein."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV126672; AAM95600.1; -.
DR InterPro; IPR003990; Pancreatiss_ac.
DR Pfam; PF00059; lectin C; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR PROSITE; PS50041; C_TYPE LECTIN 2; 1.
SQ SEQUENCE 134 AA; 14993 MW; E0E5AD9B96A53EB5 CRC64;

Query Match 35.0%; Score 307.5; DB 4; Length 134;
Best Local Similarity 55.9%; Pred. No. 2e-25;
Matches 62; Conservative 9; Mismatches 9; Indels 31; Gaps 3;

QY 1 MASRSRLRLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFYFKLRNWSDAEL---- 56
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MASRSRLRLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGFYFKLRNWSDAEYRNLL 60
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 57 -----ECQSYGNGAHLASILSLKEASTIAEYISGYQSORQPIWI 94
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 PAWPGLSRAKDQPEQ-----ISFDSGSSV---LPGHYEKPWL 97
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 6
Q8C6F9 PRELIMINARY; PRT; 126 AA.
AC Q8C6F9
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Regenerating islet-derived 1 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
   60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK075798; BAC35967.1; -.
FT NON TER 1
SQ SEQUENCE 126 AA; 14340 MW; 2564F04EA9E63094 CRC64;

Query Match 26.8%; Score 235; DB 11; Length 126;
Best Local Similarity 33.6%; Pred. No. 1.3e-17;
Matches 42; Conservative 32; Mismatches 45; Indels 6; Gaps 5;

QY 37 HKSNCYGFYFKLRNWSDAELECQSYGNGAHLASILSLKEASTIAEVI--SGYQSORQPIWI 94
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 3 YSSCYCYPTEDLTWADADLFCQNNWSG-YLVSVLSQAEQNFVASLIKESGTTDAN-VMT 60
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 95 GLHDPQKQOWIDGAMLYLRSW-SGKSMGKN-HCAEMSSNNFLTWSSNECKRQH 152
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 GLHDPKRNRRWHWSGSLFLYKSWATGSPNSSNRGCVSLTNTGYKKWKDNDCAQYSF 120
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 153 LCKYR 157
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 121 VCKFK 125
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 7
Q8PSNO PRELIMINARY; PRT; 135 AA.
AC Q8PSNO
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
```

```
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE PAL-LECTIN
OS Bitis arietans (African puff adder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Viperinae; Bitis.
OX NCBI_TaxID=8692;
RN [1]
RP MEDLINE=96161481; PubMed=8593494;
RA Nikai T., Suzuki J., Komori Y., Ohkura M., Ohizumi Y., Sugihara H.;
RT "Primary structure of the lectin from the venom of Bitis arietans
  (puff-adder).";
RL Biol. Pharm. Bull. 18:1620-1622(1995).
DR HSP; P22897; IEQG.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; Pancreat_1.
DR PRINTS; PR01504; PNCREATITAP.
DR SMART; SM00034; CLECT. 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 135 AA; 16158 MW; 3A01248FA884C5B2 CRC64;

Query Match 26.7%; Score 234; DB 13; Length 135;
Best Local Similarity 35.1%; Pred. No. 1.9e-17;
Matches 46; Conservative 17; Mismatches 64; Indels 4; Gaps 2;

QY 29 SCAPGFYHKSNCYGYFRKLRNWSDALEQSYNGAHLASILSKKEASTIAEYISGYQR 88
DB 2 NCPDMLPMNGLCYKIFDELKAWEDAERFCRKYKPGCHLASFHQYGESLEIAEYISDYHK 61

QY 89 SQ-PIWGLHDPKROQWOWIDGAMLYRWSGKS---MGNKHCAEMSSNNFLTSSN 144
DB 62 GQAEVWIGLWKKDFSWETWDRSCTDYLTDKQPDHYQKFCVELSLTGYRLWNDQ 121

QY 145 ECNKQHFELCK 155
DB 122 VCGSKNAFLCQ 132

RESULT 8
Q8JIV9 PRELIMINARY; PRT; 154 AA.
ID Q8JIV9 PRELIMINARY; PRT; 154 AA.
AC Q8JIV9;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE C-type lectin (Agglutinin-alpha 1 subunit precursor).
OS Agkistrodon acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Deinagkistrodon.
OX NCBI_TaxID=36307;
RN [1]
RP SEQUENCE FROM N.A.
RA Yu H., Xiang K., Wang Y., Liu J.;
RT "Member of C-type lectin family from Deinagkistrodon acutus.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21542539; PubMed=11686327;
RA Wang W.-J., Huang T.-F.;
RT "A novel tetrameric venom protein, agglutinin from Agkistrodon acutus,
  acts as a glycoprotein Ib agonist.";
RL Thromb. Haemost. 86:1077-1086(2001).
[3]
RP SEQUENCE FROM N.A.
RA Wang W.-J., Ling Q.-D., Huang T.-F.;
RT "Molecular structure and functional characterization of agglutinin, a
  tetrameric glycoprotein Ib-binding protein, from Formosan pit viper.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; AY091759; AAM22787.1; -.
DR EMBL; AF540645; AAN23124.1; -.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancreat_1.
DR Pfam; PF00059; lectin_c; 1.
DR PRINTS; PR01504; PNCREATITAP.
DR SMART; SM00034; CLECT. 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
KW Lectin; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 154 AGGLUCETIN-ALPHA 1 SUBUNIT.
SQ SEQUENCE 154 AA; 17317 MW; AA08E518501BECC7 CRC64;

Query Match 26.3%; Score 230.5; DB 13; Length 154;
Best Local Similarity 32.7%; Pred. No. 5.2e-17;
Matches 51; Conservative 24; Mismatches 70; Indels 11; Gaps 4;

QY 5 SMRLLLLSCLAKTGYLGIIMRSCAPGWYHKSNCYGYFRKLRNWSDALEQSYNG 64
DB 8 SFGLLVFELSCTG-ADV---DCLPGMSAYDQSCYRVFKLLKTWDDAEKCTERPKG 61

QY 65 AHLASILSKKEASTIAEYISGYRSPQIWGLHDPKROQ--QWIDGAMLYRWSGKS 122
DB 62 GHLVSIIESAGERDFAQLVSENKQTDNVMLGLKIQSGQQCSTWTDGSSVSYENFSEYQ 121

QY 123 MGNKHCHEMSSNNFLTSSNECNKQHFELCKYRP 158
DB 122 ---SKCFVLEKRTGPTWLNLCGSEYAFVCKSP 154

RESULT 9
Q9PSM4 PRELIMINARY; PRT; 135 AA.
ID Q9PSM4 PRELIMINARY; PRT; 135 AA.
AC Q9PSM4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Lectin-like protein.
OS Lachesis muta (bushmaster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Lachesis.
OX NCBI_TaxID=8752;
RN [1]
RP SEQUENCE.
RX MEDLINE=97000492; PubMed=8843577;
RA Aragon-Ortiz F., Mentale R., Auerswald E.A.;
RT "Amino acid sequence of a lectin-like protein from Lachesis muta
  stenophrys venom.";
RL Toxicon 34:763-769(1996).
DR HSP; P22897; IEQG.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancreat_1.
DR Pfam; PF00059; lectin_c; 1.
DR PRINTS; PR01504; PNCREATITAP.
DR SMART; SM00034; CLECT. 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 135 AA; 16223 MW; D5E9083A065A8F56 CRC64;

Query Match 26.1%; Score 229; DB 13; Length 135;
Best Local Similarity 34.4%; Pred. No. 6.4e-17;
Matches 45; Conservative 18; Mismatches 64; Indels 4; Gaps 2;

QY 29 SCAPGFYHKSNCYGYFRKLRNWSDALEQSYNGAHLASILSKKEASTIAEYISGYQR 88
DB 2 NCPDMLPMNGLCYKIFDELKAWEDAERFCRKYKPGCHLASFHQYGESLEIAEYISDYHK 61

QY 89 SQ-PIWGLHDPKROQWOWIDGAMLYRWSGKS---MGNKHCAEMSSNNFLTSSN 144
DB 62 GQAEVWIGLWKKDFSWETWDRSCTDYLTDKQPDHYEGKFCVELSLTGYRLWNDQ 121
```



```

QY 145 ECNKRQHFLCK 155
Db 122 VCESKNAFLCQ 132

RESULT 10
Q9CVF4
ID Q9CVF4 PRELIMINARY; PRT; 146 AA.
AC Q9CVF4 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Rat generating islet-derived, mouse homolog 3 alpha (Fragment).
GN REG3A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Small intestine;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayaishizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK008470; BAB25687.1; -.
DR HSSP; P05451; 1QDD.
DR MGD; MGI:109408; Reg3a.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancreatis_ac.
DR Pfam; PF00059; lectin_c; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
FT NON TER 1
SQ SEQUENCE 146 AA; 16161 MW; 76D73B81BECC5FB1 CRC64;

Query Match 25.9%; Score 227.5; DB 11; Length 146;
Best Local Similarity 34.5%; Pred. No. 1e-16;
Matches 48; Conservative 25; Mismatches 57; Indels 9; Gaps 4;

QY 27 RPSCAFGWFHKNCYGYFKLRNWSDALEFCQSYNGAHASILSLKEASTIAEYISG- 85
Db 8 RTSCPMGYKAYRSHSYALVMTPKSWFQADLVCCQKPSG-HLVSLSGGEASFVSSLVNGR 66

QY 86 YQRSQPIWGLHDPKQKQ-----QWIDGAMLYRSWSG--KSMGNKHCAMSSNNNF 138
Db 67 VDNVQDILWGLHDPKQKQGGGWSNSDLVNLWNGDPSSTVNRGHCGLTASSGF 126

QY 139 LTWSSNECKNRQHFLCKYR 157
Db 127 LKWGDYDCDGLPFCVKF 145

RESULT 11
Q90WI8
ID Q90WI8 PRELIMINARY; PRT; 158 AA.
AC Q90WI8 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE C-type lectin-like protein 1.
OS Bungarus fasciatus (Banded krait).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Bugarinae; Bungarus.
OX NCBI_TaxID=8613;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Zha H.-G., Zhang Y.;
RT "cDNA cloning and characterization of C-type lectin-like proteins from
RT Elapidae snakes.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF354270; AAK43584.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
KW Lectin.
SQ SEQUENCE 158 AA; 18638 MW; ECF85936FA4182B8 CRC64;

Query Match 25.9%; Score 227; DB 13; Length 158;
Best Local Similarity 33.6%; Pred. No. 1.3e-16;
Matches 44; Conservative 22; Mismatches 61; Indels 4; Gaps 2;

QY 29 SCAPGWYFKNCYGYFKLRNWSDALEFCQSYNGAHASILSLKEASTIAEYISG- 88
Db 25 TCIPIWLPKNGLCYKVFSGHKTWFDMAEYCRFKPCGCHLASLHNSADAEVFSYISDYL 84

QY 89 SQ-PIWIGLHDPKQKQWIDGAMLYRSWSGKS---MGKNKHCAMSSNNNF 144
Db 85 GQGHVWIGLRTKKYIWEWTDRTDPLPRKQKQDFHNNNEFCVEIVNFTGYLQWDD 144

QY 145 ECNKRQHFLCK 155
Db 145 NCAALRPFLCQ 155

RESULT 12
Q90WI7
ID Q90WI7 PRELIMINARY; PRT; 158 AA.
AC Q90WI7 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE C-type lectin-like protein 2.
OS Bungarus fasciatus (Banded krait).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Bugarinae; Bungarus.
OX NCBI_TaxID=8613;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Zha H.-G., Zhang Y.;
RT "cDNA cloning and characterization of C-type lectin-like proteins from
RT Elapidae snakes.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF354271; AAK43585.1; -.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancreatis_ac.
DR Pfam; PF00059; lectin_c; 1.
DR PRINTS; PR00356; ANTIFREEZEII.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.

```

```

DR PROSITE; PS00615; C_TYPE LECTIN_1; 1.
KW LECTIN; PS50041; C_TYPE LECTIN_2; 1.
SQ SEQUENCE 158 AA; 18254 MW; 5F0218970DA17453 CRC64;

Query Match 25.9%; Score 227; DB 13; Length 158;
Best Local Similarity 32.1%; Pred. No. 1.3e-16;
Matches 42; Conservative 26; Mismatches 59; Indels 4; Gaps 2;

QY 29 SCAPGWYHKSNCYGFPRKLRNWSDAELCQSYNGAHLASILSLKEASTIAEYISGYQR 88
DB 25 TCPIDWLPKNGLCYKVFSPKSLDAEMFCRFKPGCHLSIHRDADSADLAEYVSYLK 84
QY 89 SQ-PIWGLHDPKQKQOWIDGAMLYRSWSGKSMGNK---HCAEMSSNNFLTWSSN 144
DB 85 DDGNVWIGLDPKQKRTWWSDRSSNYSFNSNQEPNKNKKEYCVHLWAPTGYLKWDA 144
QY 145 ECKNRQHFLECK 155
DB 145 PCETLHPFICQ 155

RESULT 13
Q8JIV8 PRELIMINARY; PRT; 155 AA.
AC Q8JIV8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE C-type lectin.
OS Agkistrodon acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Deinagkistrodon.
OX NCBI_TaxID=36307;
RN [1]
RP SEQUENCE FROM N.A.
RA Yu H., Xiang K., Wang Y., Liu J.;
RT "Member of C-type lectin family from Deinagkistrodon acutus.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY091760; AAM22788.1; -.
DR InterPro; IPR001304; LECTIN_C.
DR InterPro; IPR003990; Pancreatins_ac.
DR Pfam; PF00059; lectin.c; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE LECTIN_2; 1.
KW LECTIN.
SQ SEQUENCE 155 AA; 17944 MW; 3E935FF53773AB94 CRC64;

Query Match 25.2%; Score 221.5; DB 13; Length 155;
Best Local Similarity 29.7%; Pred. No. 5e-16;
Matches 44; Conservative 28; Mismatches 65; Indels 11; Gaps 2;

QY 10 LLSCLAKTGVLDIIMRPSCAPGWYHKSNCYGFPRKLRNWSDAELCQSYNGAHLAS 69
DB 11 LLVFLSLGTAAD-----CPSGWSSYDGHYQVFSDLKNWDDAESFCSCQHGSRILAS 64
QY 70 ILSLKEASTIAEYISGYQRQPIWGLHDPKQKQOWIDGAMLYRSWSGKSMGNKHC 129
DB 65 IHSREBEAFVGKMSRTLYKTSMLGLNPNWKECKEWSDDTLRLDYKVWTRR-----PYC 119
QY 130 AEMSSNNFLTWSSNECKNRQHFLECKYR 157
DB 120 TMVVVKTDRIFWFRNGCEKSVSFVCKFX 147

RESULT 14
Q90W16 PRELIMINARY; PRT; 158 AA.
ID Q90W16
AC Q90W16;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

```

```

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DB C-type lectin-like protein 1.
OS Bungarus multicinctus (Many-banded krait).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Elapidae; Bungarinae; Bungarus.
OX NCBI_TaxID=8616;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Zha H.-G., Zhang Y.;
RT "cDNA cloning and characterization of C-type lectin-like proteins from Elapid snakes.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF354272; AAK43586.1; -.
DR InterPro; IPR001304; LECTIN_C.
DR InterPro; IPR003990; Pancreatins_ac.
DR Pfam; PF00059; lectin.c; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE LECTIN_2; 1.
KW LECTIN.
SQ SEQUENCE 158 AA; 18706 MW; 66B71A29D1048805 CRC64;

Query Match 24.4%; Score 214; DB 13; Length 158;
Best Local Similarity 32.1%; Pred. No. 3.3e-15;
Matches 42; Conservative 22; Mismatches 63; Indels 4; Gaps 2;

QY 29 SCAPGWYHKSNCYGFPRKLRNWSDAELCQSYNGAHLASILSLKEASTIAEYISGYQR 88
DB 25 TCPIDWLPKNGLCYKVFSPKSLDAEMFCRFKPGCHLSIHRDADSADLAEYVSYLK 84
QY 89 SQ-PIWGLHDPKQKQOWIDGAMLYRSWSGKSMGNK---HCAEMSSNNFLTWSSN 144
DB 85 DDGNVWIGLDPKQKRTWWSDRSSNYSFNSNQEPNKNKKEYCVHLWAPTGYLKWDA 144
QY 145 ECKNRQHFLECK 155
DB 145 PCETLHPFICQ 155

RESULT 15
Q9PSM8 PRELIMINARY; PRT; 125 AA.
ID Q9PSM8
AC Q9PSM8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ECLV IX/X-BP beta SUBUNIT-CA(2+)-dependent coagulation factor
DE IX/factor X-binding protein beta subunit.
OS Echin carinatus (Saw-scaled viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Viperinae; Echis.
OX NCBI_TaxID=40353;
RN [1]
RP SEQUENCE.
RX MEDLINE=96196635; PubMed=8611513;
RA Chen Y.L., Tsai I.H.;
RT "Functional and sequence characterization of coagulation factor IX/factor X-binding protein from the venom of Echis carinatus leucogaster.";
RL Biochemistry 35:5264-5271 (1996).
DR HSP; P23807; IIXX.
DR InterPro; IPR001304; LECTIN_C.
DR Pfam; PF00059; lectin.c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE LECTIN_2; 1.
SQ SEQUENCE 125 AA; 14372 MW; EDFEC2E49686CDD6 CRC64;

```

Search completed: December 31, 2003, 09:14:57
Job time : 70 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 31, 2003, 10:05:14 ; Search time 2709 Seconds
(without alignments)
201.876 Million cell updates/sec

Title: US-09-525-041-2

Perfect score: 878

Sequence: 1 MASRSNRLLLLSCLAKTGV.....LTWSNECNKRQHFLCKYRP 158

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2263443 seqs, 1730637950 residues

Total number of hits satisfying chosen parameters: 4526886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09525041/runat_31122003_091151_11986/app.query.fasta_1.327
-DB=Published Applications_NA -QFWT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-MAXLEN=0 -ALIGN=15 -MODES=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
878	100.0	474	9	US-09-922-217-1073	Sequence 1073, Ap																																							
878	100.0	474	10	US-09-833-263-1073	Sequence 1073, Ap																																							
878	100.0	474	14	US-10-035-380-1073	Sequence 1073, Ap																																							
878	100.0	477	15	US-10-100-608B-1	Sequence 1, Appli																																							
878	100.0	614	9	US-09-922-217-1075	Sequence 1075, Ap																																							
878	100.0	614	10	US-09-833-263-1075	Sequence 1075, Ap																																							
878	100.0	614	14	US-10-025-380-1075	Sequence 1075, Ap																																							
878	100.0	1114	9	US-09-922-217-1071	Sequence 1071, Ap																																							
878	100.0	1114	9	US-09-833-263-1071	Sequence 1071, Ap																																							
878	100.0	1114	10	US-09-833-263-1074	Sequence 1074, Ap																																							
878	100.0	1114	13	US-09-525-041-1	Sequence 1, Appli																																							
878	100.0	1114	14	US-10-025-380-1071	Sequence 1071, Ap																																							
878	100.0	1114	14	US-10-025-380-1074	Sequence 1074, Ap																																							
878	100.0	1130	15	US-10-157-031-113	Sequence 113, App																																							
878	100.0	1140	12	US-10-295-027-860	Sequence 860, App																																							
878	100.0	1140	12	US-10-295-027-860	Sequence 860, App																																							
878	100.0	1152	9	US-09-922-217-1072	Sequence 1072, Ap																																							
878	100.0	1152	10	US-09-833-263-1072	Sequence 1072, Ap																																							
878	100.0	1152	14	US-10-025-380-1072	Sequence 1072, Ap																																							
878	100.0	1173	9	US-09-922-217-1069	Sequence 1069, Ap																																							
878	100.0	1173	10	US-09-833-263-1069	Sequence 1069, Ap																																							
878	100.0	1173	14	US-10-025-380-1069	Sequence 1069, Ap																																							
878	100.0	1200	12	US-10-295-027-137	Sequence 137, App																																							
878	100.0	1200	12	US-10-295-027-993	Sequence 993, App																																							
878	100.0	1200	15	US-10-205-823-346	Sequence 346, App																																							
878	100.0	1309	15	US-10-106-698-2117	Sequence 2117, Ap																																							
785	89.4	551	10	US-09-962-832-3	Sequence 3, Appli																																							
785	89.4	551	13	US-09-873-367C-907	Sequence 907, App																																							
731	83.3	503	10	US-09-940-300A-1057	Sequence 1057, Ap																																							
731	83.3	503	13	US-10-033-528-1057	Sequence 1057, Ap																																							
731	83.3	503	14	US-10-033-528-1057	Sequence 1057, Ap																																							
647.5	73.7	903	13	US-09-814-353-21215	Sequence 21215, A																																							
628	71.5	481	10	US-09-998-598-25	Sequence 25, Appl																																							
485	55.2	555	10	US-09-998-598-1218	Sequence 1218, Ap																																							
458	52.2	555	9	US-09-815-343-1539	Sequence 1539, Ap																																							
443	50.5	473	10	US-09-998-598-2535	Sequence 2535, Ap																																							
430	49.0	393	14	US-09-878-178-367	Sequence 367, App																																							
430	49.0	393	14	US-10-046-935-367	Sequence 367, App																																							
430	49.0	393	15	US-10-146-502-367	Sequence 367, App																																							
429	48.9	354	10	US-09-878-178-2001	Sequence 2001, Ap																																							
429	48.9	354	14	US-10-046-935-2001	Sequence 2001, Ap																																							
429	48.9	354	15	US-10-146-502-2001	Sequence 2001, Ap																																							
429	48.9	364	10	US-09-878-178-1089	Sequence 1089, Ap																																							
429	48.9	364	14	US-10-046-935-1089	Sequence 1089, Ap																																							

ALIGNMENTS

RESULT 1

US-09-922-217-1073
; Sequence 1073, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1073

; LENGTH: 474

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-922-217-1073

Alignment Scores:

Pred. No.: 1.64e-108 Length: 474
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-525-041-2 (1-158) x US-09-922-217-1073 (1-474)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB 1 ATGGCTTCCAGAGCATGCGGCTGCTCTATTGCTGAGCTGCTGGCCAAACAGGAGTC 60
QY 21 LeuGlyAspIlelleMetArgProSerCysAlaProGlyTTPheTyrHisLysSerAsn 40
DB 61 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGGATGGTTTACCACCAAGTCCAAT 120
QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
DB 121 TGCTATGGTTACTTCAGAGAGCTGAGGAACTGGTCTGATGCCGAGCTCGAGTGTCACTCT 180
QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB 181 TACGGAAACGGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATTAGCA 240
QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTyrPheLeuHisAspProGln 100
DB 241 GAGTACATAAGTGGCTATCAGAGAGCCAGCGCATGATGGATGGCTTGCAGACCCACAG 300
QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
DB 301 AAGAGGACAGCATGGCAGTGGATGGTGGGCCATGTATCTGTACAGATCCTGGTCTGGC 360
QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
DB 361 AAGTCCATGGTGGGAAACAGCACTGTGCTGAGATGAGCTCCAATAACAACCTTTTAACT 420
QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
DB 421 TGGAGCAGCAACGAATGCAACAGCCCAACACTTCTGTGCAAGTACCGACCA 474

RESULT 2

US-09-833-263-1073
; Sequence 1073, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1073
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-263-1073

Alignment Scores:

Pred. No.: 1.64e-108 Length: 474
Score: 878.00 Matches: 158

Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-525-041-2 (1-158) x US-09-833-263-1073 (1-474)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB 1 ATGGCTTCCAGAGCATGCGGCTGCTCTATTGCTGAGCTGCTGGCCAAACAGGAGTC 60
QY 21 LeuGlyAspIlelleMetArgProSerCysAlaProGlyTTPheTyrHisLysSerAsn 40
DB 61 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGGATGGTTTACCACCAAGTCCAAT 120
QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
DB 121 TGCTATGGTTACTTCAGGAACTGGTCTGATGCCGAGCTCGAGTGTCACTCT 180
QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB 181 TACGGAAACGGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATTAGCA 240
QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTyrPheLeuHisAspProGln 100
DB 241 GAGTACATAAGTGGCTATCAGAGAGCCAGCGCATGATGGATGGCTTGCAGACCCACAG 300
QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
DB 301 AAGAGGACAGCATGGCAGTGGATGGTGGGCCATGTATCTGTACAGATCCTGGTCTGGC 360
QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
DB 361 AAGTCCATGGTGGGAAACAGCACTGTGCTGAGATGAGCTCCAATAACAACCTTTTAACT 420
QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
DB 421 TGGAGCAGCAACGAATGCAACAGCCCAACACTTCTGTGCAAGTACCGACCA 474

RESULT 3

US-10-025-380-1073
; Sequence 1073, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yudi
; APPLICANT: Smith, Carole E.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1073
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-025-380-1073

Alignment Scores:									
Pred. No.:	1.64e-108	Length:	474						
Score:	878.00	Matches:	158						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	14	Gaps:	0						
US-09-525-041-2 (1-158) x US-10-025-380-1073 (1-474)									
QY	1	MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal	20						
DB	1	ATGGCTTCAGAAGCATCGGCTGCTCTATGTGAGCTGCCCTGGCCAAAACAGGAGTC	60						
QY	21	LeuGlyAspIleleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn	40						
DB	61	CTGGGTGATATCATCATGAGACCCAGCTGTCTCTGGATGGTTTATACCAAGTCCAA	120						
QY	41	CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer	60						
DB	121	TGCTATGGTTACTTCAGGAAGCTGAGGAACCTGGTCTGATGCCGAGCTCGAGTCTCAGTCT	180						
QY	61	TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla	80						
DB	181	TACGGNAACGAGCCACCTGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA	240						
QY	81	GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln	100						
DB	241	GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATGGCTGCACGACCCACAG	300						
QY	101	LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly	120						
DB	301	AAGAGCAGCAGTGGCAGTGGATTGATGGGCCCATGTATCTGTACAGATCCTGGTCTGGC	360						
QY	121	LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr	140						
DB	361	AGTCTCATGGTGGGAACAGACCTGTCTGAGATGAGCTCCAACTTTTAACTTTTAACT	420						
QY	141	TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro	158						
DB	421	TGGAGCAGCAACGATATGCACNAGGCCCAACACTTCTGTGCAAGTACCGACCA	474						
RESULT 4									
US-10-100-608B-1									
; Sequence 1, Application US/10100608B									
; Publication No. US20030104412A1									
; GENERAL INFORMATION:									
; APPLICANT: Heiskala, Marja									
; TITLE OF INVENTION: REG-LIKE PROTEIN									
; FILE REFERENCE: CDS-261									
; CURRENT APPLICATION NUMBER: US/10/100,608B									
; CURRENT FILING DATE: 2002-09-10									
; PRIOR APPLICATION NUMBER: 60/276,414									
; PRIOR FILING DATE: 2002-03-16									
; NUMBER OF SEQ ID NOS: 45									
; SOFTWARE: PatentIn version 3.1									
; SEQ ID NO 1									
; LENGTH: 477									
; TYPE: DNA									
; ORGANISM: Human									
US-10-100-608B-1									
Alignment Scores:									
Pred. No.:	1.65e-108	Length:	477						
Score:	878.00	Matches:	158						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	15	Gaps:	0						
US-09-525-041-2 (1-158) x US-10-100-608B-1 (1-477)									
QY	1	MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal	20						

Db 136 ATGGCTTCAGAGCATGGCGTCTCTATTTGCTGAGCTGCCTGCCCAACAGGAGTC 195
Qy 21 LeuGlyAspIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
Db 196 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTCTGATGGTTTACCAAGTCCAA 255
Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db 256 TGCTATGGTTACTTCAGGAAGCTGAGGAACCTGGTCTGATCGGAGCTCGAGTGTCACT 315
Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 316 TACGGAACCGGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 375
Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 376 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCTTCGACGCCACAG 435
Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db 436 AAGAGCCAGCAGTGGCAGTGGATTGATGGGCCCATGTATCTGTACAGATCTCTGTCTGC 495
Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnPheLeuThr 140
Db 496 AAGTCCATGGTGGGAACAGCAGCTGTCTGAGATGAGCTCCATATACAACTTTTAACT 555
Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db 556 TGGAGCAGCAACGAATGCAACAGCGCCAACTTCTGTGCAAGTACCGACCA 609

RESULT 6

US-09-833-263-1075
; Sequence 1075, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolck, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1075
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-263-1075

Alignment Scores:
Pred. No.: 2,37e-108 Length: 614
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-525-041-2 (1-158) x US-09-833-263-1075 (1-614)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db 136 ATGGCTTCAGAGCATGGCGTCTCTATTTGCTGAGCTGCCTGCCCAACAGGAGTC 195
Qy 21 LeuGlyAspIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
Db 196 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGATGGTTTACCAAGTCCAA 255
Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db 256 TGCTATGGTTACTTCAGGAAGCTGAGGAACCTGGTCTGATCGGAGCTCGAGTGTCACT 315

Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 316 TACGGAACCGGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 375
Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 376 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCTTCGACGCCACAG 435
Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db 436 AAGAGCCAGCAGTGGCAGTGGATTGATGGGCCCATGTATCTGTACAGATCTCTGTCTGC 495
Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnPheLeuThr 140
Db 496 AAGTCCATGGTGGGAACAGCAGCTGTCTGAGATGAGCTCCATATACAACTTTTAACT 555
Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db 556 TGGAGCAGCAACGAATGCAACAGCGCCAACTTCTGTGCAAGTACCGACCA 609

RESULT 7

US-10-025-380-1075
; Sequence 1075, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolck, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: Wang, Gordon E.
; APPLICANT: King, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Panger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1075
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-025-380-1075

Alignment Scores:
Pred. No.: 2,37e-108 Length: 614
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-525-041-2 (1-158) x US-10-025-380-1075 (1-614)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db 136 ATGGCTTCAGAGCATGGCGTCTCTATTTGCTGAGCTGCCTGCCCAACAGGAGTC 195
Qy 21 LeuGlyAspIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
Db 196 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGATGGTTTACCAACAGTCCAA 255

QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
DB 256 TGCATATGTTACTTCAGGAAGCTGAGAACTGGTCTGATCCGAGCTCGAGTGTCACTCT 315
QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB 316 TAGGAAACCGAGCCACCTGGCATCTATCTGAGTTTAAAGAGCCAGCACCATAGCA 375
QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
DB 376 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCTCGACGACCCACAG 435
QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
DB 436 AAGAGGCGAGCTGGCAGTGGATTGATGGGGCCATGATCTGTACAGATCTCTGGTCTGGC 495
QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
DB 496 AAGTCCATGGGTGGAAACAAGCACTGTGCTGAGATGAGCTCCCAATACCAACTTTTAACT 555
QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
DB 556 TGGAGCAGCAACGAATGCAACAAGCGCCACACTTCTGTGCAAGTACCGACCA 609

RESULT 8

US-09-922-217-1071
; Sequence 1071, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-217-1071

Alignment Scores:
Pred. No.: 5.6e-108 Length: 1114
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-525-041-2 (1-158) x US-09-922-217-1071 (1-1114)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB 111 ATGGCTTCCAGAGCATGCGGCTGCTCTATTGCTGAGCTGCTGCCAAAACAGGAGTC 170
QY 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
DB 171 CTGGGTGATATCATGAGACCCAGCTGTGCTCCGGATGGTTTACCAAGTCCAAT 230
QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60

DB 231 TGCATATGTTACTTCAGGAAGCTGAGAACTGGTCTGATCCGAGCTCGAGTGTCACTCT 290
QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB 291 TAGGAAACCGAGCCACCTGGCATCTATCTGAGTTTAAAGAGCCAGCACCATAGCA 350
QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
DB 351 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCTCGACGACCCACAG 410
QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
DB 411 AAGAGGCGAGCTGGCAGTGGATTGATGGGGCCATGATCTGTACAGATCTCTGGTCTGGC 470
QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
DB 471 AAGTCCATGGGTGGAAACAAGCACTGTGCTGAGATGAGCTCCCAATACCAACTTTTAACT 530
QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
DB 531 TGGAGCAGCAACGAATGCAACAAGCGCCACACTTCTGTGCAAGTACCGACCA 584

RESULT 9

US-09-922-217-1074
; Sequence 1074, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1074
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-217-1074

Alignment Scores:
Pred. No.: 5.6e-108 Length: 1114
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-525-041-2 (1-158) x US-09-922-217-1074 (1-1114)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB 111 ATGGCTTCCAGAGCATGCGGCTGCTCTATTGCTGAGCTGCTGCCAAAACAGGAGTC 170
QY 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
DB 171 CTGGGTGATATCATGAGACCCAGCTGTGCTCCGGATGGTTTACCAAGTCCAAT 230
QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60

```
Db 231 TGCTATGGTTACTTCAGGAAGCTGAGGAAGCTGGTCTGATGCCGAGCTCGAGTGTCACTCT 290
Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 291 TACGGAACGAGGCCCTGGCATCTATCTCTGAGTTAAAGGAAGCCAGCACCATTAGCA 350
Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 351 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGAATTGGCTTGCACGCCACAG 410
Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db 411 AAGAGGCAGCAGTGGCATGGAATTCATGGGCCCATGTATCTGTACAGATCTCTGGTCTGGC 470
Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
Db 471 AAGTCCATGGTGGGAACAGCACTGTCTGAGATGAGCTCCAATAACAACCTTTTAACT 530
Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db 531 TGGAGCAGCAAGAAATGCAACAAGCCCAACACTTCTGTGCAAGTACCACCA 584
RESULT 10
US-09-833-263-1071
; Sequence 1071, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1071
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-263-1071
Alignment Scores:
Pred. No.: 5.6e-108 Length: 1114
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-525-041-2 (1-158) x US-09-833-263-1071 (1-1114)
Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db 111 ATGGCTTCCAGAGCATCGGCTGCTCTATGTGTGAGTGCCTGGCCAAACAGGAGTC 170
Qy 21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
Db 171 CTGGGTGATATCATCATGAGACCAGCTGTGCTCTGGATGGTTTACCACAACTCCAAT 230
Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db 231 TGCTATGGTTACTTCAGGAAGCTGAGAACTGGTCTGATGCCGAGCTCGAGTGTCACTCT 290
Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 291 TACGGAACGAGGCCCTGGCATCTATCTCTGAGTTAAAGGAAGCCAGCACCATTAGCA 350
Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 351 GAGTACATAAGTGGCTATCAGGAAGCTGAGAACTGGTCTGATGCCGAGCTCGAGTGTCACTCT 290
Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 291 TACGGAACGAGGCCCTGGCATCTATCTCTGAGTTAAAGGAAGCCAGCACCATTAGCA 350
Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 351 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGAATTGGCTTGCACGCCACAG 410
```

```
Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db 411 AAGAGGCAGCAGTGGCATGGAATTCATGGGCCCATGTATCTGTACAGATCTCTGGTCTGGC 470
Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
Db 471 AAGTCCATGGTGGGAACAGCACTGTCTGAGATGAGCTCCAATAACAACCTTTTAACT 530
Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db 531 TGGAGCAGCAAGAAATGCAACAAGCCCAACACTTCTGTGCAAGTACCACCA 584
RESULT 11
US-09-833-263-1074
; Sequence 1074, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1074
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-263-1074
Alignment Scores:
Pred. No.: 5.6e-108 Length: 1114
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-525-041-2 (1-158) x US-09-833-263-1074 (1-1114)
Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db 111 ATGGCTTCCAGAGCATCGGCTGCTCTATGTGTGAGTGCCTGGCCAAACAGGAGTC 170
Qy 21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
Db 171 CTGGGTGATATCATCATGAGACCAGCTGTGCTCTGGATGGTTTACCACAACTCCAAT 230
Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db 231 TGCTATGGTTACTTCAGGAAGCTGAGAACTGGTCTGATGCCGAGCTCGAGTGTCACTCT 290
Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 291 TACGGAACGAGGCCCTGGCATCTATCTCTGAGTTAAAGGAAGCCAGCACCATTAGCA 350
Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 351 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGAATTGGCTTGCACGCCACAG 410
Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db 411 AAGAGGCAGCAGTGGCATGGAATTCATGGGCCCATGTATCTGTACAGATCTCTGGTCTGGC 470
Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
Db 471 AAGTCCATGGTGGGAACAGCACTGTCTGAGATGAGCTCCAATAACAACCTTTTAACT 530
```


RESULT 14

US-10-025-380-1074
; Sequence 1074, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secretist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1074
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-025-380-1074

Alignment Scores:

Pred. No.: 5,6e-108 Length: 1114
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-525-041-2 (1-158) x US-10-025-380-1074 (1-1114)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db 111 ATGGCTTCCAGAACATCGGCTGCTCTTATTTGCTGAGCTGCCTGGCCAAACAGGAGTC 170
Qy 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
Db 171 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGGATGGTTTTTACCACAGTCCAAT 230
Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db 231 TGCTATGGTTACTTCCAGAACCTCAGGAACCTGGTCTGATGCCGAGCTCGAGTGTCACT 290
Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 291 TACGGAAACCGAGCCCACTCGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 350
Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 351 GAGTACATAAGTGGCTATCAGAGAACCCAGCCGATGATGGATTGGCCCTGCAGACCCACAG 410
Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db 411 AAGAGGCAGCAGTGGCAGTGGATTGATGGGGCCATGATCTCTGTACAGATCTCTGTCTGGC 470
Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
Db 471 AAGTCCATGGGTGGGACCAAGCACTGTGCTGAGATGAGCTCCCAATAACAACTTTTAACT 530

Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db 531 TGGAGCAGCAACGAATGCAACAGCGCCACACTTCTCTGTGCAAGTACCGACCA 584

RESULT 15

US-10-157-031-113
; Sequence 113, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 113
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-157-031-113

Alignment Scores:

Pred. No.: 5,71e-108 Length: 1130
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-525-041-2 (1-158) x US-10-157-031-113 (1-1130)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db 123 ATGGCTTCCAGAACATCGGCTGCTCTTATTTGCTGAGCTGCCTGGCCAAACAGGAGTC 182
Qy 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
Db 183 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGGATGGTTTTTACCACAGTCCAAT 242
Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db 243 TGCTATGGTTACTTCCAGGAAGCTGAGAACTGGTCTGATGCCGAGCTCGAGTGTCACT 302
Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 303 TACGGAAACCGAGCCCACTCGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 362
Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 363 GAGTACATAAGTGGCTATCAGAGAACCCAGCCGATGATGGATTGGCCCTGCAGCACCCACAG 422
Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db 423 AAGAGGCAGCAGTGGCAGTGGATTGATGGGGCCATGATCTCTGTACAGATCTCTGTCTGGC 482
Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
Db 483 AAGTCCATGGGTGGGACCAAGCACTGTGCTGAGATGAGCTCCCAATAACAACTTTTAACT 542
Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db 543 TGGAGCAGCAACGAATGCAACAGCGCCACACTTCTCTGTGCAAGTACCGACCA 596

Search completed: December 31, 2003, 12:21:55

Job time : 2712 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 31, 2003, 09:15:58 ; Search time 251 Seconds
(without alignments)
1699.248 Million cell updates/sec

Title: US-09-525-041-2
Perfect score: 878
Sequence: 1 MASRMRLLLLLSCLAKTV.....LTWSSNECKRQHFLCKYRP 158

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-DB=/cn2.1/USPTO spool_p/US09525041/runat_31122003_091148_11855/app.query.fasta_1.327
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09525041 @CGN 1 1 490 @runat_31122003_091148_11855 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03:.*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	878	100.0	474	19	AAV29036	Open reading frame
2	878	100.0	474	22	AAI29508	Human secretory si
3	878	100.0	474	22	AAF82117	Human TSA7005 prot
4	878	100.0	474	25	ABZ33694	Human colon specif
5	878	100.0	477	24	ABZ33694	Human colon specif
6	878	100.0	477	24	ABZ33694	Human REG-like pro
7	878	100.0	477	24	ABZ33694	Human REG-like pro
8	878	100.0	614	19	AAV29156	cDNA encoding a hu
9	878	100.0	614	22	AAI29510	Human Reg 1-gamma
10	878	100.0	1060	22	ABZ33696	Human colon specif
11	878	100.0	1060	22	ABZ33696	Human TSA7005 prot
12	878	100.0	1114	18	AAT51784	Human colon specif
13	878	100.0	1114	20	AAK03195	cDNA sequence of a
14	878	100.0	1114	21	AAAG2951	Human colon specif
15	878	100.0	1114	22	AAI29506	Human colon specif
16	878	100.0	1114	22	AAI29509	Human colon specif
17	878	100.0	1114	25	ABZ33692	Human colon specif
18	878	100.0	1114	25	ABZ33695	Human colon specif
19	878	100.0	1152	19	AAV29035	Human protein comp
20	878	100.0	1152	22	AAI29507	Human secretory si
21	878	100.0	1152	25	ABZ33693	Human colon specif
22	878	100.0	1170	22	AAH57432	Human intestine ce
23	878	100.0	1173	22	AAI29505	CSI-152 clone full
24	878	100.0	1173	25	ABZ33691	Human colon tumour
25	878	100.0	1200	24	ABV72313	Nucleotide sequenc
26	878	100.0	1200	24	ABQ60776	Human Reg IV cDNA
27	878	100.0	1255	21	AAH09227	Human cancer speci
28	878	100.0	1309	22	AAH35025	Human colon cancer
29	878	100.0	1613	23	ABV23028	Human prostate exp
30	878	100.0	1613	23	ABV23055	Human prostate exp
31	878	100.0	1613	23	ABV28864	Human prostate exp
32	865	98.5	1372	23	ABV28891	Human prostate exp
33	785	89.4	551	24	ABL62570	Cancer specific ge
34	785	89.4	551	24	ABL62570	Colon adenocarcino
35	740	84.3	497	24	ABL67552	Oesophagus cancer
36	731	83.3	503	24	ABK27780	Human colon cancer
37	709	80.8	508	25	ABZ36760	cDNA encoding colo
38	628	71.5	481	24	ABV86714	Human GENSET codin
39	577	65.7	611	24	ABK39550	Human colon cancer
40	485	55.2	555	24	ABV87907	Colon adenocarcino
41	478	54.4	527	24	ABK29631	Human colon cancer
42	462	52.6	539	24	ABK29641	Colon adenocarcino
43	458	52.2	555	23	AAS58863	Colon adenocarcino
44	443	50.5	473	24	ABV89224	cDNA #1539 encodin
45	436	49.7	426	23	ABV05145	Human colon cancer
						Human prostate exp

ALIGNMENTS

RESULT 1
AAV29036
ID AAV29036 standard; cDNA; 474 BP.
XX
AC AAV29036;
XX
DT 10-AUG-1998 (first entry)
XX
DE Open reading frame human protein comprising secretory signal 3.
XX
KW Human protein; secretory signal; nutritional source; cytokine;
KW immunity; haematopoiesis; activin; inhibitor; tumour; chemotactic;
KW chemokinetic; thrombolytic; anti-inflammatory; inhibition;
KW stomach cancer cell; ds.
OS Homo sapiens.
XX

FH Key Location/Qualifiers
 FT CDS 1..474
 FT /*tag= a
 FT /product= "human protein comprising secretory
 FT signal"
 XX
 XX WO9811217-A2.
 XX
 XX 19-MAR-1998.
 XX
 XX 12-SEP-1997; 97WO-JP03239.
 XX
 XX 13-SEP-1996; 96JP-0243060.
 XX
 XX (PROT-) PROTEGENE INC.
 XX (SAGA) SAGAMI CHEM RES CENTRE.
 XX
 XX Kato S, Kobayashi M, Sekine S, Yamaguchi T;
 XX
 XX WPI; 1998-207380/18.
 XX P-PSDB; AAW37866.
 XX
 XX Human proteins with secretory signal sequences - used to treat
 XX immune deficiencies, infections, tumours, and haematopoietic
 XX disorders, etc.
 XX
 XX Claim 3; Page 82; 131pp; English.
 XX
 XX This is the nucleotide sequence of the open reading frame of a novel
 XX human protein comprising a secretory signal (AAV29035), isolated from
 XX stomach cancer cells. Its proteins can be used as nutritional sources
 XX or supplements. The proteins may also have cytokine functions.
 XX CC immune modulating functions, haematopoiesis regulating activity,
 XX CC activin/inhibin regulating activity, chemotactic/chemokinetic activity,
 XX CC haemostatic and thrombolytic activity, receptor/ligand activity,
 XX CC anti-inflammatory activity, tumour inhibition activity.
 XX
 XX Sequence 474 BP; 126 A; 116 C; 129 G; 103 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 1.14e-86 Length: 474
 Score: 878.00 Matches: 158
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 19 Gaps: 0
 US-09-525-041-2 (1-158) x AAV29036 (1-474)
 QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 DB 1 ATGCTTCCAGAGCATGCGCTCTCTATGTCTGAGCTGCCCTGGCCAAACAGGAGTC 60
 QY 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 DB 61 CTGGGTGATATCATCATGAGACCCAGCTGTCTCTCTGGATGGTTTACCAAGTCCAAAT 120
 QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 DB 121 TGCTATGTTTACTTTCAGGAAGCTGAGAACTGGTCTGTATGCCGAGCTCGAGTGTCTCT 180
 QY 61 TyrGlyAanGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 DB 181 TACGGAACGGAGCCACCTGGCATCTATCTCTGATGTTAAAGGAGCCAGCACCATTAGCA 240
 QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 DB 241 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGATGGTGGCTGCAGCACCACAG 300
 QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 DB 301 AAGAGGAGCAGTGGCAGTGGATTGATGGGGCCATGTATCTGTACAGATCTCTGTCTGGC 360

QY 121 LysSerMetGlyCysAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
 DB 361 AAGTCCATGGTGGGAACAAGCACCTGTCTGAGTCCCAATAACAACATTTTAACT 420
 QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
 DB 421 TGGAGCAGCAACGAATGCAACCAAGGCCCAACACTTCTCTGTGCAAGTACCGACCA 474
 RESULT 2
 AAI29508
 ID AAI29508 standard; cDNA; 474 BP.
 XX
 AC AAI29508;
 DT 12-OCT-2001 (first entry)
 XX
 XX Human secretory signal nucleotide sequence 3 cDNA sequence #2.
 DE Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 XX immunogenic; gene therapy; vaccine; colonic cancer; ss.
 KW Homo sapiens.
 XX
 OS WO200149716-A2.
 XX
 PN 12-JUL-2001.
 XX
 PD 29-DEC-2000; 2000WO-US35596.
 XX
 PF 30-DEC-1999; 99US-0476296.
 XX
 PR 10-JAN-2000; 2000US-0480321.
 PR 15-FEB-2000; 2000US-0504629.
 PR 06-MAR-2000; 2000US-0519444.
 PR 19-MAY-2000; 2000US-0575251.
 PR 29-JUN-2000; 2000US-0609448.
 PR 28-AUG-2000; 2000US-0649811.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;
 PI King GE, Wang T, Jiang Y;
 XX
 XX WPI; 2001-441847/47.
 DR
 XX Colon tumor associated proteins and nucleic acids useful for the
 PT prevention, diagnosis and treatment of colonic cancer -
 XX
 XX Claim 2; Page 465; 472pp; English.
 XX
 CC The present invention describes colon tumor associated proteins (I) and
 CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
 CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
 CC (II) may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate colon tumor associated protein (TCAP)
 CC expression, such as colonic cancer. For example, (I) and (II) may be
 CC used to treat disorders associated with decreased expression by
 CC rectifying mutations or deletions in a patient's genome that affect the
 CC activity of TCAPs by expressing inactive proteins or to supplement the
 CC patients own production of them. Additionally, (II) may be used to
 CC produce the TCAP proteins, by inserting the nucleic acids into a host
 CC cell culturing the cell to express the protein. (II) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC polynucleotide chain reaction (PCR) and hybridisation assays to detect and
 CC quantitate the presence of similar nucleic acids in samples, and
 CC therefore which patients may be in need of restorative therapy. (I) may
 CC also be used as antigens in the production of antibodies against TCAPs
 CC and in assays to identify modulators of TCAP expression and activity.
 CC Anti-(I) antibodies and antagonists may also be used to down regulate
 CC TCAP expression and activity. The anti-(I) antibodies may also be used
 CC as diagnostic agents for detecting the presence of TCAPs in samples
 CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AAI29460 to AAI29512
 CC and AAM24494 to AAM24523 represent nucleotide and amino acid sequences
 CC given in the exemplification of the present invention.

PR	15-JUL-1999;	99JP-0201279.
XX	(SAKA) OTSUKA PHARM CO LTD.	
XX	WPI; 2001-303742/32.	
XX	P-PsDB; AAB74934.	
PT	TSA7005 gene, encoding a polypeptide useful for the diagnosis and treatment of diseases associated with its expression -	
XX	Claim 3; Page 23; 25pp; Japanese.	
CC	The present sequence encodes a human TSA7005 protein which shares 32% homology with human and mouse Reg proteins, and 34% homology with the rat Reg protein. TSA7005 has pancreatic beta cell growth activity and hypoglycaemic activity. The TSA7005 protein can be used for the CC diagnosis and treatment of diseases associated with the gene and its CC expression product.	
XX	Sequence 474 BP; 126 A; 116 C; 129 G; 103 T; 0 other;	
Alignment Scores:		
Pred. No.:	1..14e-86	Length: 474
Score:	878.00	Matches: 158
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	22	Gaps: 0
US-09-525-041-2 (1-158) x AAF82117 (1-474)		
QY	1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20	
Db	1 ATGGCTTCCAGAACATCGGGCTGCTCTATTGTGAGCTGCCTGCCCAAAACAGGAGTC 60	
QY	21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTyrlHisLysSerAsn 40	
Db	61 CTGGTGATATCATCATGAGACCAGCTGTCTCTGGATGGTTTACCAACAGTCCAAT 120	
QY	41 CysTyrgLyTyrrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuCysGlnSer 60	
Db	61 CTGGTGATATCATCATGAGACCAGCTGTCTCTGGATGGTTTACCAACAGTCCAAT 120	
QY	121 TGCTATGTTACTTTCAGGAAGCTGAGAACCTGGTCTGATGCCGAGTCCAGTCTTAACT 180	
Db	61 TyrgLyAsnGlyAlaHisLeuAlaSerlleLeuSerLeuLysGluAlaSerThrIleAla 80	
Db	181 TAGCGAAACGGACCCACCTGGCATCTATCTCTGAGTTTAAGAAGCACGACCATAGCA 240	
QY	81 GluTyrlleSerGlyTyrgLnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100	
Db	241 GAGTACATAAGTGGCTTATCAGAGAAGCCAGCCGATATGGATTGGCTGCAGCACCCACAG 300	
QY	101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrrLeuTyrrArgSerTrpSerGly 120	
Db	301 RAGAGGACGAGTGGGAGTGGATTGATGGGCGCATGTATCTGTACAGATCCTGGTCTGGC 360	
QY	121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140	
Db	361 AAGTCCATGGTGGGAACAAGCACTGTCTGATGAGTCCCAATAACAACCTTTTAACT 420	
QY	141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrrArgPro 158	
Db	421 TGGAGCAGCAACGAATGCAACAGCCCAACACTTCTCTGCAAGTAGTACCGACCA 474	
RESULT 3		
AAF82117		
ID	AAF82117 standard; cDNA; 474 BP.	
AC	AAF82117;	
XX	27-JUN-2001 (first entry)	
DE	Human TSA7005 protein encoding cDNA SEQ ID NO:2.	
XX	Human; TSA7005; Reg; pancreatic beta cell growth; hypoglycaemic; diagnosis; ss.	
KW	Homo sapiens.	
OS		
XX	Key Location/Qualifiers	
FT	1..474	
CDS	/tag= a	
FT	/partial	
FT	/product= "TSA7005 protein"	
FT	/note= "no stop codon given"	
XX	JP2001025389-A.	
PN		
PD	30-JAN-2001.	
XX	15-JUL-1999; 99JP-0201279.	
PF		
XX		

KW Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;
 KW tumour; immune response; immunostimulant; cytostatic; vaccine;
 KW gene; ss.
 XX Homo sapiens.
 XX WO200283070-A2.
 XX 24-OCT-2002.
 XX 09-APR-2002; 2002WO-US11475.
 XX 10-APR-2001; 2001US-0833263.
 PR 03-AUG-2001; 2001US-0922217.
 PR 19-DEC-2001; 2001US-0025380.
 XX (CORI-) CORIXA CORP.
 XX Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;
 PI Wang T, Jiang Y, Smith CL, King GB, Wang A, Clapper JD;
 PI Skeiky YAW, Fanger GR, Vedvick TS, Carter D;
 XX WPI; 2003-067548/06.
 DR New polynucleotide, useful for the preparation of a composition for
 XX stimulating an immune response against, or treating, cancer.
 XX Disclosure; Page 466-467; 537pp; English.
 XX The present invention describes compounds (I) for the immunotherapy and
 CC diagnosis of colon cancer. Also described: (1) a method for detecting
 CC the presence of cancer in a patient; (2) a method for stimulating and/or
 CC expanding T cells specific for a tumour protein; (3) an isolated T cell
 CC population comprising T cells prepared by the method of (2); (4) a method
 CC for stimulating an immune response in a patient; (5) a method for
 CC treating cancer in a patient; and (6) a method for inhibiting the
 CC development of cancer in a patient. (I) have immunostimulant and
 CC cytostatic activities and can be used in vaccines. ABZ332646 to ABZ33725
 CC and ABP55343 to ABP55391 represent human colon cancer/tumour related
 CC sequences used in the exemplification of the present invention.
 XX SQ Sequence 474 BP; 126 A; 116 C; 129 G; 103 T; 0 other;

Alignment Scores:
 Pred. No.: 1.14e-86 Length: 474
 Score: 878.00 Matches: 158
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 25 Gaps: 0

US-09-525-041-2 (1-158) x ABZ33694 (1-474)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 Db 1 ATGGCTTCCAGAACATGGCGTCTCTCTATTGTGAGTGGCTGGCCAAACAGGAGTC 60
 Qy 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 Db 61 CTGGGTGATATCATCATGAGACCCAGCTGCTCTCTGGATGGTTTACCAAGTCCCAAT 120
 Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 Db 121 TGCTATGGTTACTTACGAGAGCTGAGAACTGGTCTGATGCCGAGCTCGAGTTCAGTCT 180
 Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 Db 181 TACGGAACGGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATTAGCA 240
 Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 Db 241 GAGTACATAAGTGGCTATCAGAGAAGCCGCGATATGGATGGCTCGCAGCCCAACAG 300

Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 Db 301 AAGAGGCAGCAGTGGCAGTGGATTGATGGGGCCATGTATCTGTACAGATCCTGGTCTGGC 360
 Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
 Db 361 AAGTCCATGGTGGGAACCAAGCACTGTGTGTGATGAGCTCCCAATAACAACTTTTAACT 420
 Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
 Db 421 TGGAGCAGCAACGATGCAACAGCCCAACACTTCTGTGCAAGTACCGACCA 474
 RESULT 5
 AAD47239
 ID AAD47239 standard; cDNA; 477 BP.
 XX AC AAD47239;
 XX DT 24-FEB-2003 (first entry)
 XX DE Human REG-like protein (RELp) cDNA.
 XX KW Human; REG-like protein; RELp; tumour; cancer; therapy; chromosome 1;
 XX gene; ss.
 XX OS Homo sapiens.
 XX PH Key Location/Qualifiers
 FT CDS 1..477
 FT /*tag= a
 FT /product= "Human REG-like protein"
 FT sig_peptide 1..66
 FT /*tag= b
 FT mat_peptide 67..474
 FT /*tag= c
 FT /product= "Human mature REG-like protein"
 XX EP1241269-A2.
 XX 18-SEP-2002.
 XX 15-MAR-2002; 2002EP-0251876.
 XX 16-MAR-2001; 2001US-276414P.
 XX (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
 XX Heiskala M;
 XX WPI; 2002-684095/74.
 XX P-PSDB; AAE29829.
 XX Detecting the presence of a tumor comprises detecting the concentration
 of a Reg Like Protein or the presence or quantity of a nucleic acid
 encoding it
 XX Example 1; Page 13; 26pp; English.
 XX The invention relates to a method for detecting REG-like protein (RELp)
 and its nucleic acid sequence. The method is useful for detecting the
 presence of a tumour. Kits comprising an antibody specific for RELp and
 reagents for detecting the antibody, or a nucleic acid complementary to
 a portion of a nucleic acid encoding RELp, are useful for identifying
 the presence of cancer, characterise the cancer, or monitor the course
 of treatment of cancer. The present sequence is human RELp cDNA used to
 CC illustrate the method of the invention. Human RELp gene is located at
 CC chromosome 1.
 XX SQ Sequence 477 BP; 127 A; 116 C; 130 G; 104 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.14e-86 Length: 477
 Score: 878.00 Matches: 158

AAV29156 standard; cDNA; 614 BP.
 AAV29156;
 21-AUG-1998 (first entry)
 cDNA encoding a human Reg I-gamma protein.
 Human; Reg I-gamma protein; C-type lectin; Incyte clone 1310334;
 regulation; cell growth; development; tumorigenesis; neurodegeneration;
 inhibition; treatment; prevention; neoplasia; metastasis;
 neurodegenerative change; Alzheimer's disease; Down's syndrome;
 regeneration; pancreatic beta-cells; diabetes; ss.
 Homo sapiens.
 Key Location/Qualifiers
 CDS 136..612
 /*tag= a
 WO9816640-A1.
 23-APR-1998.
 07-OCT-1997; 97WO-US18174.
 11-OCT-1996; 96US-0729103.
 (INCY-) INCYTE PHARM INC.
 Bandman O, Goli SK;
 WPI; 1998-251287/22.
 P-PSDB; AAW37929.
 New isolated human Reg I-gamma protein - useful for developing
 products for treating, e.g. diabetes, tumours or neuro-degenerative
 disease such as Alzheimer's
 Claim 4; Fig 1; 72pp; English.
 The present sequence encodes a human Reg I-gamma protein, which
 comprises a C-type lectin. The sequence was identified in Incyte clone
 1310334. Reg I-gamma protein is involved in regulation of cell growth
 and development. Since the overexpression of reg proteins is associated
 with tumorigenesis and neurodegeneration, inhibition of human Reg
 I-gamma expression can be used for treating or preventing neoplasia or
 metastasis and neurodegenerative changes associated with Alzheimer's
 disease and other disorders of the central nervous system, e.g. Down's
 syndrome. Reg I-gamma can also be used in therapeutics to induce
 regeneration of pancreatic beta-cells in the treatment of diabetes. The
 products can also be used for detection for, e.g. expression of REG
 I-gamma, diagnosis and drug screen
 Sequence 614 BP; 168 A; 143 C; 171 G; 132 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.6e-86 Length: 614
 Score: 878.00 Matches: 158
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 19 Gaps: 0
 US-09-525-041-2 (1-158) x AAV29156 (1-614)
 Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 Db 136 ATGGCTTCCAGAGCATGCGCTGCTCTATTGCTGAGCTGCTGGCCAAACAGGAGTC 195
 Qy 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 Db 196 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGGTGGTGGTTTACCACCAAGTCCAA 255

Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTyrSerAspAlaGluLeuGluCysGlnSer 60
 Db 256 TGCTATGGTTACTTTCAGGAAGCTGAGAACTGGTCTGATGCCGAGCTCAGTGTGAGTCT 315
 Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 Db 316 TACGGAACCGAGCCACCTGGCATCTATCTCTAGTTTAAAGAGCCAGCACCATTAGCA 375
 Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 Db 376 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCTGCACGCCACAG 435
 Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 Db 436 AAGAGCAGCAGTGGCAGTGGATTGATGGGGCATGTATCTGTACAGATCTCTGGTCTGGC 495
 Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
 Db 496 AAGTCCATGGTGGGAACAAGCACCTGTCTGAGATGAGCTCCAATAACAACATTTTAACT 555
 Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
 Db 556 TGGAGCAGCAACGAATGCAACAGCCCAACACTTCTCTGTGCAAGTACCGACCA 609
 RESULT 8
 AAI29510
 ID AAI29510 standard; cDNA; 614 BP.
 XX
 AC AAI29510;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Human Reg 1-gamma protein cDNA sequence.
 XX
 KW Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 KW immunogenic; gene therapy; vaccine; colonic cancer; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200149716-A2.
 XX
 PD 12-JUL-2001.
 XX
 PF 29-DEC-2000; 2000WO-US35596.
 XX
 PR 30-DEC-1999; 99US-0476296.
 PR 10-JAN-2000; 2000US-0480321.
 PR 15-FEB-2000; 2000US-0504629.
 PR 06-MAR-2000; 2000US-0519444.
 PR 19-MAY-2000; 2000US-0575251.
 PR 29-JUN-2000; 2000US-0609448.
 PR 28-AUG-2000; 2000US-0649811.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Lodes MJ, Secret H, Benson DR, Meagher MJ, Scolk JA;
 PI King GE, Wang T, Jiang Y;
 XX
 DR WPI; 2001-441847/47.
 XX
 PT Colon tumor associated proteins and nucleic acids useful for the
 PT prevention, diagnosis and treatment of colonic cancer -
 XX
 PS Claim 2; Page 465; 472pp; English.
 XX
 CC The present invention describes colon tumour associated proteins (I) and
 CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
 CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
 CC (II) may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate colon tumour associated protein (TCAP)
 CC expression, such as colonic cancer. For example, (I) and (II) may be
 CC used to treat disorders associated with decreased expression by

CC rectifying mutations or deletions in a patient's genome that affect the
 CC activity of TCAPs by expressing inactive proteins or to supplement the
 CC patients own production of them. Additionally, (II) may be used to
 CC produce the TCAP proteins, by inserting the nucleic acids into a host
 CC cell culturing the cell to express the protein. (II) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC polymerase chain reaction (PCR) and hybridisation assays to detect and
 CC quantitate the presence of similar nucleic acids in samples, and
 CC therefore which patients may be in need of restorative therapy. (I) may
 CC also be used as antigens in the production of antibodies against TCAPs
 CC and in assays to identify modulators of TCAP expression and activity.
 CC Anti-(I) antibodies and antagonists may also be used to down regulate
 CC TCAP expression and activity. The anti-(I) antibodies may also be used
 CC as diagnostic agents for detecting the presence of TCAPs in samples
 CC (e.g. By enzyme linked immunosorbent assay (ELISA)). AAI28460 to AAI29512
 CC and AAM2494 to AAM24523 represent nucleotide and amino acid sequences
 CC given in the exemplification of the present invention.

XX
 SQ Sequence 614 BP; 168 A; 143 C; 171 G; 132 T; 0 other;

Alignment Scores:

Pred. No.: 1.6e-86 Length: 614
 Score: 878.00 Matches: 158
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-525-041-2 (1-158) x AAI29510 (1-614)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 DB 136 ATGGCTTCAGAAAGCATGCGGCTGCTCTATTGCTGAGCTGCTGCCAAAACAGGAGTC 195
 QY 21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 DB 196 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGGATGCTTTTACCAAGTCCAA 255
 QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 DB 256 TGCTATGTTACTTCAGGAGCTGAGGAAGCTGCTCTGATGCCAGCTCGAGTGTCTAGTCT 315
 QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 DB 316 TACGGAACCGAGCCAGCCATCTCTCTGAGTTTAAAGGAAGCCAGCACCACCATAGCA 375
 QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTyrIleGlyLeuHisAspProGln 100
 DB 376 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATGGATGGCTGCGACGCCACGAC 435
 QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 DB 436 AAGAGCAGCAGTGGCAGTGAGTATGATGGGCGCATGTATCTGTACAGATCTCTGCTGGC 495
 QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
 DB 496 AAGTCCATGGGTGGGAACAAGCAGCTGTCTGAGATGAGCTCCCAATAACAACCTTTTAACT 555
 QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
 DB 556 TGGAGCAGCAACGAATGCAACAAGCCCAACACTTCTCTGTGAAGTACCGACCA 609

RESULT 9

ABZ33696

ID ABZ33696 standard; cDNA; 614 BP.

XX AC ABZ33696;

XX 30-JAN-2003 (first entry)

XX Human colon specific gene cDNA sequence SEQ ID NO:1075.

XX Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;

KW tumour; immune response; immunostimulant; cytostatic; vaccine;
 KW gene; ss.

XX Homo sapiens.

PN WO200283070-A2.

XX 24-OCT-2002.

XX 09-APR-2002; 2002WO-US11475.

XX 10-APR-2001; 2001US-0833263.

PR 03-AUG-2001; 2001US-0922217.

PR 19-DEC-2001; 2001US-0025380.

PA (CORI-) CORIXA CORP.

PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
 PI Wang T, Jiang Y, Smith CL, King GE, Wang A, Clapper JD;
 PI Skeiky YAW, Fanger GR, Vedvick TS, Carter D;

XX WPI; 2003-067548/06.

XX New polynucleotide, useful for the preparation of a composition for
 PT stimulating an immune response against, or treating, cancer -

XX Disclosure; Page 467; 537pp; English.

XX The present invention describes compounds (I) for the immunotherapy and
 CC diagnosis of colon cancer. Also described: (1) a method for detecting
 CC the presence of cancer in a patient; (2) a method for stimulating and/or
 CC expanding T cells specific for a tumour protein; (3) an isolated T cell
 CC population comprising T cells prepared by the method of (2); (4) a method
 CC for stimulating an immune response in a patient; (5) a method for
 CC treating cancer in a patient; and (6) a method for inhibiting the
 CC development of cancer in a patient. (I) have immunostimulant and
 CC cytostatic activities and can be used in vaccines. ABZ32646 to ABZ33725
 CC and ABP55343 to ABP55391 represent human colon cancer/tumour related
 CC sequences used in the exemplification of the present invention.

XX Sequence 614 BP; 168 A; 143 C; 171 G; 132 T; 0 other;

Alignment Scores:

Pred. No.: 1.6e-86 Length: 614
 Score: 878.00 Matches: 158
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 25 Gaps: 0

US-09-525-041-2 (1-158) x ABZ33696 (1-614)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 DB 136 ATGGCTTCAGAAAGCATGCGGCTGCTCTATTGCTGAGCTGCTGCCAAAACAGGAGTC 195
 QY 21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 DB 196 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGATGGTGTTCACCAAGTCCAA 255
 QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 DB 256 TGCTATGTTACTTCAGGAGCTGAGGAAGCTGCTCTGATGCCAGCTCGAGTGTCTAGTCT 315
 QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 DB 316 TACGGAACCGAGCCAGCCATCTCTCTGAGTTTAAAGGAAGCCAGCACCACCATAGCA 375
 QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTyrIleGlyLeuHisAspProGln 100
 DB 376 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATGGATGGCTGCGACGCCACGAC 435
 QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120

```

Db      436 AAGAGGAGCAGTGGCAGTGGATTGATGGGCCATGATCTGTACAGATCCTGGTCTGGC 495
QY      121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnPheLeuThr 140
Db      496 AAGTCCATGGTGGGAACAAGCACTGTGCTGAGATGAGCTCCAATAACAACCTTTTAACT 555
QY      141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db      556 TGGAGCAGCAACGATGCAACAGGCCACACCTTCTGTGCAAGTACCGCCA 609

RESULT 10
AAF82118
ID      AAF82118 standard; cDNA; 1060 BP.
XX
AC      AAF82118;
XX
DT      27-JUN-2001 (first entry)
XX
DE      Human TSA7005 protein encoding cDNA SEQ ID NO:3.
XX
KW      Human; TSA7005; Reg; pancreatic beta cell growth; hypoglycaemic;
XX      diagnosis; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
FT      Location/Qualifiers
FT      67..543
FT      /*tag= a
FT      /product= "TSA7005 protein"
XX
PN      JP2001025389-A.
XX
PD      30-JAN-2001.
XX
PF      15-JUL-1999; 99JP-0201279.
XX
PR      15-JUL-1999; 99JP-0201279.
XX
PA      (SAKA ) OTSUKA PHARM CO LTD.
XX
DR      WPI; 2001-303742/32.
XX
DR      P-PSDB; AAB74934.
XX
PT      TSA7005 gene, encoding a polypeptide useful for the diagnosis and
XX      treatment of diseases associated with its expression -
XX
PS      Disclosure; Page 23-24; 25pp; Japanese.
XX
CC      The present sequence encodes a human TSA7005 protein which shares
XX      32% homology with human and mouse Reg proteins, and 34% homology with
XX      the rat Reg protein. TSA7005 has pancreatic beta cell growth activity
XX      and hypoglycaemic activity. The TSA7005 protein can be used for the
XX      diagnosis and treatment of diseases associated with the gene and its
XX      expression product.
XX
SQ      Sequence 1060 BP; 269 A; 278 C; 247 G; 264 T; 2 other;

Alignment Scores:
Pred. No.: 3.33e-86 Length: 1060
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-525-041-2 (1-158) x AAF82118 (1-1060)

QY      1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db      67 ATGGCTTCCAGAACATCGCGCTGCTCTATGCTGAGCTGCGCCAAACACAGGAGTC 126
QY      21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40

```

```

Db      127 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTCTGGATGGTTTTACCACAGTCCAAT 186
QY      41 CyeTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db      187 TGTCTATGGTTACTTTCAGGAAGCTGAGAACTGTCTGTATGCCGAGCTCGAGTGTAGTCT 246
QY      61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db      247 TACGGAACCGAGGCCACCTGGCATCTATCTCTGAGTTTAAAGGAGCCAGCACCATAGCA 306
QY      81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db      307 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCTGCACGCCACAG 366
QY      101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db      367 AAGAGCAGCAGTGGCAGTGGATTGATGGGGCATGTATCTGTACAGATCCTGGTCTGGC 426
QY      121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
Db      427 AAGTCCATGGTGGGAACAAGCACCTGTGCTGAGATGAGCTCCAATAACAACCTTTTAACT 486
QY      141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db      487 TGGAGCAGCAACGATGCAACAGGCCAACACCTTCTGTGCAAGTACCGACCA 540

RESULT 11
AAT51784
ID      AAT51784 standard; cDNA; 1114 BP.
XX
AC      AAT51784;
XX
DT      31-MAY-1997 (first entry)
XX
DE      Human colon specific protein cDNA.
XX
KW      Colon specific gene; colon cancer; metastasis; diagnosis;
XX      gene therapy; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
FT      Location/Qualifiers
FT      111..587
FT      /*tag= a
FT      /product= mature colon specific protein
FT      complement (111..130)
FT      primer_bind
FT      /*tag= b
FT      /note= "5' primer site for PCR amplification
FT      and baculovirus expression"
FT      complement (112..128)
FT      primer_bind
FT      /*tag= c
FT      /note= "5' primer site for PCR amplification
FT      and bacterial expression"
FT      primer_bind
FT      575..592
FT      /*tag= d
FT      /note= "3' primer site for PCR amplification and
FT      baculovirus or bacterial expression"
XX
XX      WO9639541-A1.
XX
XX      12-DEC-1996.
XX
XX      06-JUN-1995; 95WO-US07169.
XX
XX      06-JUN-1995; 95WO-US07169.
XX      (HUMA-) HUMAN GENOME SCI INC.
XX
XX      Dillon RJ, Li Y, Soppet DR;
XX      WPI; 1997-043162/04.
XX      P-PSDB; AAW12691.
XX

```

XX	New isolated colon specific gene - used to develop prods. for use in
PT	the diagnosis and treatment of colon disorders, partic. colon
PT	cancer.
XX	
PS	Claim 4; Page 51-52; 64pp; English.
XX	
CC	A cDNA clone (AAT51784) codes for a colon specific protein (AAW12691)
CC	that may be useful as a diagnostic marker for colon cancer. It is
CC	believed that the presence of active transcription of the colon
CC	specific gene in non-colon cells of a host is indicative of colon
CC	cancer metastases. The cDNA, deposited as ATCC 97129, was isolated
CC	from a human colon cancer cDNA library. It can be used to design
CC	diagnostic probes, to produce recombinant colon specific protein in
CC	transformed host cells, in the gene therapy of patients having need
CC	of a colon specific gene protein, or to design antisense constructs
CC	useful for treatment of colon cancer.
XX	
SQ	Sequence 1114 BP; 288 A; 291 C; 262 G; 273 T; 0 other;
	Alignment Scores:
	Pred. No.: 3.56e-86 Length: 1114
	Score: 878.00 Matches: 158
	Percent Similarity: 100.00% Conservatives: 0
	Best Local Similarity: 100.00% Mismatches: 0
	Query Match: 100.00% Indels: 0
	DB: 18 Gaps: 0
	US-09-525-041-2 (1-158) x AAT51784 (1-1114)
QY	1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal
Db	111 ATGGCTTCAGAAAGCATGCGCTGCTCTATTGCTGAGCTGCCTGAGCAACAGGAGT
QY	21 LeuGlyAspIlelleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn
Db	171 CTGGGTGATATCATCATGAGACCCAGCTGTCTCTGGATGGTTTTTACCACAAAGTCCA
QY	41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer
Db	231 TGCTATGGTTACTTCAGAAAGCTGAGGNACTGGTCTGATGCCAGCTCGAGTGTCA
QY	61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla
Db	291 TACGGAAACGGAGCCACCTGGCATCTATCTCTGAGTTTTAAAGGAAGCCAGCACCAT
QY	81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln
Db	351 GAGTACATAAAGTGCTATCATGAGAAGCCAGCCGATATGGATTGGCTGCACGACCC
QY	101 LysArgGlnGlnTrpGlnTTPileAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly
Db	411 AAGAGCGACGAGTGGCAGTGGATTGATGGGCCCATGATCTGTGTACAGATCTCTGT
QY	121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr
Db	471 AAGTCCATGGTGGGAACAAGCACTGTGCTGAGATGAGCTCCCAATPAACAACCTTTT
QY	141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db	531 TGGAGCAGCAACGAAATGCAACAAGCGCCAACTTCCTCTGTGCAAGTACCGACCA 584
	RESULT 12
	AA03195
	ID AAX03195 standard; cDNA; 1114 BP.
	XX
	AA03195;
	XX
	25-MAR-1999 (first entry)
	DT
	XX
	cDNA sequence of a human colon specific gene.
	DE
	XX
	Human; colon specific gene; diagnosis; colon disorder; colon cancer;
	XX

QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 DB 411 AAGAGGCGAGCAGTGGCAGTGATTGATGGGCGCATGTATCTGTACAGATCCTGGTCTGGC 470

QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
 DB 471 AAGTCCATGGGTGGGAACAAGCACTGTGCTGAGATGAGCTCCAAATAACAACCTTTTAACT 530

QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
 DB 531 TGGAGCAGCAGCAAGTGCACACAGCGCCACACTTCTGTGTGCAAGTACCGACCA 584

RESULT 13
 AAA62951
 ID AAA62951 standard; cDNA; 1114 BP.
 XX
 AC AAA62951;
 XX
 DT 14-NOV-2000 (first entry)
 XX
 DE Human colon specific protein cDNA sequence.
 XX
 KW Human; colon specific; colon cancer; metastasis; diagnose; treatment;
 KW cytostatic; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 111..587
 FT /*tag= a
 FT /product= "Colon specific protein"

XX US6080722-A.
 XX 27-JUN-2000.
 XX 29-SEP-1998; 98US-0162508.
 XX 06-JUN-1995; 95US-0468413.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Dillon PJ, Soppet DR, Li Y;
 XX WPI: 2000-464055/40.
 XX P-PSDB; AAB12900.
 XX
 XX Novel human colon specific polypeptides and polynucleotides for
 XX diagnosis and treatment of colon cancer, for screening compounds which
 XX interact with polypeptide, for synthesis of DNA and manufacture of DNA
 XX vectors -
 XX
 XX Disclosure; Fig 1; 20pp; English.
 XX
 XX This invention relates to a purified human protein, which is primarily
 XX expressed in tissue derived from the colon. The protein is 152 amino
 XX acids in length and exhibits cytostatic activity. The present sequence
 XX represents cDNA encoding the colon specific protein. The protein can be
 XX used in the diagnosis and treatment of colon cancer, and it is thought
 XX that abnormally high levels of the gene expression in non-colon cells is
 XX an indication of colon cancer metastasis.
 XX
 XX Sequence 1114 BP; 288 A; 291 C; 262 G; 273 T; 0 other;
 SQ

Alignment Scores:
 Pred. No.: 3.56e-86 Length: 1114
 Score: 878.00 Matches: 158
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0

US-09-525-041-2 (1-158) x AAA62951 (1-1114)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 DB 111 ATGGCTTCCAGAAGCATCGGCTGCTCTTATTTGCTGAGCTGCTGGCCAAAACAGGAGTC 170

QY 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 DB 171 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGGATGGTTTTTACCACCAAGTCCAAT 230

QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 DB 231 TGTATGGTTACTTTCAGGAAGCTGAGGAACCTGCTGTGATGCCGAGCTCGAGTGTCACTCT 290

QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 DB 291 TACGGAACGGAGCCACCTGGCATCTATCTGAGTTTAAAGAGGAGCCAGCACCATAGCA 350

QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 DB 351 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGATGGCTGCACGACCCACAG 410

QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 DB 411 AAGAGCAGCAGTGGCAGTGGATTGATGGGGCCATGTATCTGTACAGATCCTGGTCTGGC 470

QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
 DB 471 AAGTCCATGGGTGGGAACAAGCACTGTGCTGAGATGAGCTCCAAATAACAACCTTTTAACT 530

QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
 DB 531 TGGAGCAGCAGCAAGTGCACACAGCGCCACACTTCTGTGTGCAAGTACCGACCA 584

RESULT 14
 AAI29506
 ID AAI29506 standard; cDNA; 1114 BP.
 XX
 AC AAI29506;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Human colon specific gene cDNA sequence #1.
 XX
 KW Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 KW immunogenic; gene therapy; vaccine; colonic cancer; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200149716-A2.
 XX
 XX 12-JUL-2001.
 XX
 XX 29-DEC-2000; 2000WO-US35596.
 XX
 XX 30-DEC-1999; 99US-0476296.
 XX 10-JAN-2000; 2000US-0480321.
 XX 15-FEB-2000; 2000US-0504629.
 XX 06-MAR-2000; 2000US-0519444.
 XX 19-MAY-2000; 2000US-0575251.
 XX 29-JUN-2000; 2000US-0609448.
 XX 28-AUG-2000; 2000US-0649811.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 XX Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
 XX King GE, Wang T, Jiang Y;
 XX WPI; 2001-441847/47.
 XX
 XX Colon tumor associated proteins and nucleic acids useful for the
 XX prevention, diagnosis and treatment of colonic cancer -
 XX
 XX Claim 2; Page 464; 472pp; English.
 PS

XX The present invention describes colon tumour associated proteins (I) and
CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be
CC used to treat disorders associated with decreased expression by
CC rectifying mutations or deletions in a patient's genome that affect the
CC activity of TCAPs by expressing inactive proteins or to supplement the
CC patients own production of them. Additionally, (II) may be used to
CC produce the TCAP proteins, by inserting the nucleic acids into a host
CC cell culturing the cell to express the protein. (II) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC polymerase chain reaction (PCR) and hybridisation assays to detect and
CC quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. (I) may
CC also be used as antigens in the production of antibodies against TCAPs
CC and in assays to identify modulators of TCAP expression and activity.
CC Anti-(I) antibodies and antagonists may also be used to down regulate
CC TCAP expression and activity. The anti-(I) antibodies may also be used
CC as diagnostic agents for detecting the presence of TCAPs in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)). AAI28460 to AAI29512
CC and AAM24494 to AAM24523 represent nucleotide and amino acid sequences
CC given in the exemplification of the present invention.
XX
SQ Sequence 1114 BP; 288 A; 291 C; 262 G; 273 T; 0 other;

Alignment Scores:
Pred. No.: 3,568-86 Length: 1114
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-525-041-2 (1-158) x AAI29506 (1-1114)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB 111 ATGGCTTCAGAGCATGCGGTCTCTTATTTGCTGAGCTGCTGCCAANAACAGAGTTC 170
QY 21 LeuGlyAspIleMetArgProSerCysAlaProGlyTrpPheTrpHisLysSerAsn 40
DB 171 CTGGGTGATCATCATGACCCAGCTGTCTCTGGATGTTTACCAAGTCCAAAT 230
QY 41 CysTrpGlyTrpPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
DB 231 TGTCTATGGTTACTTCAGGAAGCTGAGGAACCTGGTCTGATGCCGAGCTCGAGTCT 290
QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB 291 TACGGANACGGAGCCCACTTGGCACTATCTCTGAGTTTAAAGAACCCAGCACATAGCA 350
QY 81 GlyTrpIleSerGlyTrpGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
DB 351 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGATTTGGCTGCAACCCACAG 410
QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTrpLeuTrpArgSerTrpSerGly 120
DB 411 AAGAGGCACACATGGCAGTGGATTTGATGGGGCCATGATCTCTACAGATCTCTGGTCTGC 470
QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
DB 471 AAGTCCATGGTGGGACAGCAGCTGTCTGATGAGTGTCCAAATACCACTTTTAACT 530
QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTrpArgPro 158
DB 531 TGGAGCAGCAGCATGATGCAACAGCCCAACTCTCTGTGCAAGTACCGACCA 584
RESULT 15
AAI29509
ID AAI29509 standard; cdna; 1114 BP.

XX AAI29509;
AC 12-OCT-2001 (first entry)
DT Human colon specific gene cDNA sequence #2.
DE Human; immunotherapy; diagnosis; colon cancer; colon tumour;
KW immunogenic; gene therapy; vaccine; colonic cancer; ss.
XX Homo sapiens.
XX WO200149716-A2.
XX 12-JUL-2001.
XX 29-DEC-2000; 2000WO-US35596.
XX 30-DEC-1999; 99US-0476296.
XX 10-JAN-2000; 2000US-0480321.
XX 15-FEB-2000; 2000US-0504629.
XX 06-MAR-2000; 2000US-0519444.
XX 19-MAY-2000; 2000US-0575251.
XX 29-JUN-2000; 2000US-0609448.
XX 28-AUG-2000; 2000US-0649811.
XX (CORI-) CORIXA CORP.
XX Xu J, Lodes MJ, Secret H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;
XX WPI; 2001-441847/47.
XX Colon tumor associated proteins and nucleic acids useful for the
PT prevention, diagnosis and treatment of colonic cancer -
PT
XX
XX Claim 2; Page 465; 472pp; English.

The present invention describes colon tumour associated proteins (I) and the polynucleotides (II) that encode them. (I) have cytostatic activity. (I) and (II) can be used in gene therapy and vaccine production. (I) and (II) may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate colon tumour associated protein (TCAP) expression, such as colonic cancer. For example, (I) and (II) may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of TCAPs by expressing inactive proteins or to supplement the patients own production of them. Additionally, (II) may be used to produce the TCAP proteins, by inserting the nucleic acids into a host cell culturing the cell to express the protein. (II) and its complementary sequences may also be used as DNA probes in diagnostic polymerase chain reaction (PCR) and hybridisation assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. (I) may also be used as antigens in the production of antibodies against TCAPs and in assays to identify modulators of TCAP expression and activity. Anti-(I) antibodies and antagonists may also be used to down regulate TCAP expression and activity. The anti-(I) antibodies may also be used as diagnostic agents for detecting the presence of TCAPs in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). AAI28460 to AAI29512 and AAM24494 to AAM24523 represent nucleotide and amino acid sequences given in the exemplification of the present invention.

SQ Sequence 1114 BP; 288 A; 291 C; 262 G; 273 T; 0 other;

Alignment Scores:
Pred. No.: 3,568-86 Length: 1114
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-525-041-2 (1-158) x AA129509 (1-1114)

```
Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db 111 ATGGCTTCCAGAGCATGGGCTGCTCTATTGTGTAGCTGCTGGCCAAAACAGGAGTC 170
Qy 21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
Db 171 CTGGGTGATATCATCATGAGACCCAGCTGTCTCTGGATGGTGTATACCAAGTCCCAAT 230
Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db 231 TGCTATGGTTACTTCAGGAAGCTGAGGAACCTGGTCTGATGCCGAGCTCGAGTGTCACTCT 290
Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 291 TACGGAACGGAGCCCACTGGCACTCATCTGTAGTTTAAAGGAGCCAGCACCATAGCA 350
Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 351 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCTTGACGCCACACAG 410
Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db 411 AAGAGGCAGCAGTGGCAGTGGATTGATGGGCCCATGTATCTGTACAGATCCTGTCTTGGC 470
Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
Db 471 AAGTCCATGGGTGGAAACAGCACTGTGTGAGATGAGCTCCCAATAACAACCTTTTAACT 530
Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db 531 TGGAGCAGCAACGAATGCACAAGCGCCCAACACTTCTCTGTGCAAGTACCGACCA 584
```

Search completed: December 31, 2003, 09:29:07
Job time : 254 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 31, 2003, 09:16:53 ; Search time 2146 Seconds
(without alignments)
3011.986 Million cell updates/sec

Title: US-09-525-041-2
Perfect score: 878
Sequence: 1 MASRMRLLLLSLAKTGV.....LTWSSNECNKROHFLCKYRP 158

Scoring table:
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 288711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-O=/cpn2_1/USPTO_epool_p/US09525041/runat_31122003_091149_11864/app_query.fasta_1.327
-DB=GenEmbl -QFMT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09525041 @CGN 1 1 4958 @runat_31122003_091149_11864 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	gb.ba.*
2:	gb.htg.*
3:	gb.in.*
4:	gb.om.*
5:	gb.ov.*
6:	gb.pat.*
7:	gb.ph.*
8:	gb.pl.*
9:	gb.pr.*
10:	gb.ro.*
11:	gb.sta.*
12:	gb.sv.*
13:	gb.un.*
14:	gb.vi.*
15:	em.ba.*
16:	em.fun.*
17:	em.hum.*
18:	em.in.*
19:	em.mu.*
20:	em.om.*
21:	em.or.*
22:	em.ov.*
23:	em.pat.*
24:	em.ph.*
25:	em.pl.*
26:	em.ro.*
27:	em.sta.*
28:	em.un.*

ALIGNMENTS

RESULT 1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	878	100.0	474	6	AX193506 Sequence
2	878	100.0	474	6	BD062766 Human pro
3	878	100.0	474	6	E52141 TSA7005 gen
4	878	100.0	477	6	AX537651 Sequence
5	878	100.0	614	6	AR058965 Sequence
6	878	100.0	614	6	AX193508 Sequence
7	878	100.0	1060	6	E52142 TSA7005 gen
8	878	100.0	1114	6	AR030953 Sequence
9	878	100.0	1114	6	AR100599 Sequence
10	878	100.0	1114	6	AX193504 Sequence
11	878	100.0	1114	6	AX193507 Sequence
12	878	100.0	1130	6	AX676920 Sequence
13	878	100.0	1130	9	AY007243 Homo sapi
14	878	100.0	1152	6	AX193505 Sequence
15	878	100.0	1152	6	BD062775 Human pro
16	878	100.0	1161	9	AF345934 Homo sapi
17	878	100.0	1173	6	AX193502 Sequence
18	878	100.0	1192	9	BC017089 Homo sapi
19	878	100.0	1200	6	BD167401 Method of
20	878	100.0	1200	9	AF254415 Homo sapi
21	878	100.0	1518	9	AY126670 Homo sapi
22	785	89.4	551	6	AX330398 Sequence
23	785	89.4	551	6	AX335380 Sequence
24	740	84.3	497	6	AX351470 Sequence
25	731	83.3	503	6	AX396842 Sequence
26	709	80.8	508	6	AX588758 Sequence
27	709	80.8	508	6	BD108868 EST and e
28	600.5	68.4	1017	10	BC019465 Mus muscu
29	577	65.7	611	6	AX379034 Sequence
30	544	62.0	683	9	AY126671 Homo sapi
31	478	54.4	527	6	AX379115 Sequence
32	462	52.6	539	6	AX379125 Sequence
33	458	52.2	555	6	AX261888 Sequence
34	430	49.0	396	6	AX340120 Sequence
35	429	48.9	357	6	AX341754 Sequence
36	429	48.9	364	6	AX351280 Sequence
37	429	48.9	367	6	AX340842 Sequence
38	429	48.9	378	6	AX192921 Sequence
39	429	48.9	381	6	AX340101 Sequence
40	429	48.9	381	6	AX340869 Sequence
41	429	48.9	381	6	AX341040 Sequence
42	417	47.5	374	6	AX341657 Sequence
43	416	47.4	555	6	AX378982 Sequence
44	406	46.2	382	6	AX340142 Sequence
45	395	45.0	350	6	AX261862 Sequence

```
AX193506
LOCUS AX193506 474 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 1073 from Patent WO0149716.
ACCESSION AX193506
VERSION AX193506.1 GI:15211446
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,
Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
JOURNAL Patent: WO 0149716-A 1073 12-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source
BASE COUNT 126 a 116 c 129 g 103 t
ORIGIN
Alignment Scores:
Pred. No.: 3.53e-83 Length: 474
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-525-041-2 (1-158) x AX193506 (1-474)
Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db 1 ATGGCTCCAGAACGATCGGCTGCTCTCTATTGTGAGCTGCTGGCCAAACAGAGATC 60
Qy 21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
Db 61 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTCTGATGGTGTATACCAACAATCCAAT 120
Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db 121 TGCATGGTTACTTCAGAAAGCTGAGAACTGGTCTGATGCCGAGCTCGAGTGTCACT 180
Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 181 TACGGAAACGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 240
Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 241 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGATGGCTGCGACGCCACAG 300
Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db 301 AAGAGGACAGCAGTGGCAGTGGATTTGATGGGCCATGATCTGTACAGATCCTGTCTGGC 360
Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
Db 361 AAGTCCATGGTGGGAACAGCACTGTGCTGAGATGAGTCCATACCAACTTTTAACT 420
Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db 421 TGGAGCAGCAACGAATGCAACAGGCCAACACTTCTCTGTGCAAGTACCGACCA 474
RESULT 2
BD062766 474 bp DNA linear PAT 27-AUG-2002
LOCUS BD062766
DEFINITION Human proteins having secretory signal sequences and DNAs encoding
these proteins.
ACCESSION BD062766
BD062766.1 GI:22608369
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 474)
AUTHORS Kato,S., Sekine,S., Kimura,T. and Kobayashi,M.
TITLE Human proteins having secretory signal sequences and DNAs encoding
these proteins
JOURNAL Patent: JP 2001506484-A 3 22-MAY-2001;
SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC
COMMENT OS Homo sapiens (human)
PN JP 2001506484-A/3
PD 22-MAY-2001
PF 12-SEP-1997 JP 1998513509
PR 13-SEP-1996 JP 8/243060
PI SEISHI KATO,SHINGO SEKINE,TOMOKO KIMURA,MIDORI KOBAYASHI PC
C12N15/12,C07K14/47,C12N15/62
CC Strandedness: Double;
CC Topology: Linear;
PH Key Location/Qualifiers.
FEATURES
source
1..474
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 126 a 116 c 129 g 103 t
ORIGIN
Alignment Scores:
Pred. No.: 3.53e-83 Length: 474
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-525-041-2 (1-158) x BD062766 (1-474)
Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db 1 ATGGCTCCAGAACGATCGGCTGCTCTCTATTGTGAGCTGCTGGCCAAACAGAGATC 60
Qy 21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
Db 61 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTCTGATGGTGTATACCAACAATCCAAT 120
Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db 121 TGCATGGTTACTTCAGAAAGCTGAGAACTGGTCTGATGCCGAGCTCGAGTGTCACT 180
Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 181 TACGGAAACGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 240
Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 241 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGATGGCTGCGACGCCACAG 300
Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db 301 AAGAGGACAGCAGTGGCAGTGGATTTGATGGGCCATGATCTGTACAGATCCTGTCTGGC 360
Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
Db 361 AAGTCCATGGTGGGAACAGCACTGTGCTGAGATGAGTCCATACCAACTTTTAACT 420
Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db 421 TGGAGCAGCAACGAATGCAACAGGCCAACACTTCTCTGTGCAAGTACCGACCA 474
RESULT 3
```

```

E52141
LOCUS       E52141 TSA7005 gene.
DEFINITION E52141
ACCESSION  E52141
VERSION    E52141.1 GI:18629624
KEYWORDS   JP 2001025389-A/1.
SOURCE     unidentified
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 474)
AUTHORS    Ogawara,T., Suzuki,M. and Ozaki,K.
TITLE      TSA7005 Gene
JOURNAL    Patent: JP 2001025389-A 1 30-JAN-2001;
COMMENT    OTSUKA PHARMACEUT CO LTD
PN         JP 2001025389-A/1
PD         30-JAN-2001
PF         15-JUL-1999 JP 1999201279
PR
PI         TSUYOSHI OGAWARA,MIKIO SUZUKI,KOICHI OZAKI
PC         C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10//A61K31/00,
PC         A61K38/00,A61K48/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02 CC

FH Key      Location/Qualifiers
FT source   1..474
FT          /organism='Unknown'

FEATURES             Location/Qualifiers
source              1..474
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

BASE COUNT      126 a 116 c 129 g 103 t

Alignment Scores:
Pred. No.:      3,53e-83      Length:      474
Score:          878.00      Matches:    158
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:    0
DB:             6      Gaps:      0

US-09-525-041-2 (1-158) x E52141 (1-474)

QY      1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB      1 ATGGCTTCAGAGAGATGGCGCTGCTCTATGCTGAGCTGCTGGCCAAAACAGGAGTC 60

QY      21 LeuGlyAspIleileMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
DB      61 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGGATGGTTTACCAAGTCCAAT 120

QY      41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
DB      121 TGCTATGGTTACTTCAGGAAGCTGAGAACTGGTCTGATGCCGAGCTCGAGTCTCAGTCT 180

QY      61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB      181 TACGGAACCGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 240

QY      81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
DB      241 GAGTACATAGTGGCTATCAGAGAAGCCAGCCGATATGATGGCTGCACGACCCACAG 300

QY      101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
DB      301 AAGAGCAGCAGTGGCAGTGGATTTGATGGGCCCATGTATCTGTACAGATCTCTGGTCTGGC 360

QY      121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
DB      361 AAGTCCATGGGTGGGAACCAAGCACCTGTGTGATGAGCTCCCAATAACAACTTTTAACT 420

QY      141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
DB      421 TGGAGCAGCAAGATGCAACAGGCCCAACACTTCTCTGTGCNAGTACCGACCA 474

RESULT 4
AX537651
LOCUS       AX537651
DEFINITION Sequence 1 from Patent EP1241269.
ACCESSION  AX537651
VERSION    AX537651.1 GI:25269611
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Heiskala,M.
TITLE      Method for detecting reg-like protein and nucleic acids coding therefor
JOURNAL    Patent: EP 1241269-A 1 18-SEP-2002;
COMMENT    Ortho-Clinical Diagnostics, Inc. (US)
FEATURES             Location/Qualifiers
source              1..477
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

BASE COUNT      127 a 116 c 130 g 104 t

Alignment Scores:
Pred. No.:      3,55e-83      Length:      477
Score:          878.00      Matches:    158
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:    0
DB:             6      Gaps:      0

US-09-525-041-2 (1-158) x AX537651 (1-477)

QY      1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
DB      1 ATGGCTTCAGAGAGATGGCGCTGCTCTATGCTGAGCTGCTGGCCAAAACAGGAGTC 60

QY      21 LeuGlyAspIleileMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
DB      61 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGGATGGTTTACCAAGTCCAAT 120

QY      41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
DB      121 TGCTATGGTTACTTCAGGAAGCTGAGAACTGGTCTGATGCCGAGCTCGAGTCTCAGTCT 180

QY      61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
DB      181 TACGGAACCGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 240

QY      81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
DB      241 GAGTACATAGTGGCTATCAGAGAAGCCAGCCGATATGATGGCTGCACGACCCACAG 300

QY      101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
DB      301 AAGAGCAGCAGTGGCAGTGGATTTGATGGGCCCATGTATCTGTACAGATCTCTGGTCTGGC 360

QY      121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
DB      361 AAGTCCATGGGTGGGAACCAAGCACCTGTGTGATGAGCTCCCAATAACAACTTTTAACT 420

QY      141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
DB      421 TGGAGCAGCAAGATGCAACAGGCCCAACACTTCTCTGTGCNAGTACCGACCA 474

RESULT 5
AR058965

```


PF 15-JUL-1999 JP 1999201279
PR
PI TSUYOSHI OGAWARA,MIKIO SUZUKI,KOICHI OZAKI
PC C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10/A61K31/00.
PC A61K38/00,A61K48/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02 CC

FH Key Location/Qualifiers
FT CDS (67) . . (540) .
FEATURES
source Location/Qualifiers
1..1060
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 269 a 278 c 247 g 264 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 8,548-83 Length: 1060
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-525-041-2 (1-158) x E52142 (1-1060)

QY	1	MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal	20
Db	67	ATGGCTTCAGAAAGCATGGGCTGCTCTATTGTAGCTGCTGGCCAAACAGAGATC	126
QY	21	LeuGlyAspIleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn	40
Db	127	CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTCTGGATGGTTTATACCAAGTCCAAT	186
QY	41	CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer	60
Db	187	TGCTATGGTTACTTCAGAAAGCTGAGGAACCTGGTCTGATGCCAGGCTCGAGTGTCACTCT	246
QY	61	TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla	80
Db	247	TACGGAAACGGAGCCACCTGGCATCTATCTCTAGTTTAAAGNAGCCAGCACCATAGCA	306
QY	81	GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln	100
Db	307	GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCTTGACAGCCACAG	366
QY	101	LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly	120
Db	367	NAAGGCGACGACGTGGCAGTGGATTGATGGGCGCATGATCTGTACAGATCCTGGTCTGGC	426
QY	121	LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr	140
Db	427	AGTTCATGGGTGGGAACAGACATGCTCGATGAGTCCCAATACAACTTTTAACT	486
QY	141	TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro	158
Db	487	TGGAGCAGCAACGAATGCAACAAGCGCCACACTTCTGTGCAAGTACCGACCA	540

RESULT 8
AR030953
LOCUS
DEFINITION Sequence 1 from patent US 5861494.
ACCESSION AR030953
VERSION AR030953.1 GI:5944167
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1114)
AUTHORS Soppet,D.R., Li,Y. and Dillon,P.J.
TITLE Colon specific gene and protein
JOURNAL Patent: US 5861494-A 1 19-JAN-1999;

FEATURES	source	Location/Qualifiers
BASE COUNT	288 a 291 c 262 g 273 t	
ORIGIN		1..1114 /organism="unknown"
Alignment Scores:		
Pred. No.:	9.01e-83	Length: 1114
Score:	878.00	Matches: 158
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	6	Gaps: 0
US-09-525-041-2 (1-158) x AR030953 (1-1114)		
Qy	1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaIysThrGlyVal	20
Db	111 ATGGCTTCCAGAACATCGCGCTGCTCTCTATTGTGTAGCTGCTGCGCCAAACAGGAGTC	170
Qy	21 LeuGlyAspIleIleMetArgProSerCysAlaProGlyTTPheTyRHisIysSerAsn	40
Db	171 CTGGGTGATATCATCATGAGACCCAGCTGTGCTCTGGATGGTTTTACCACAAAGTCCAA	230
Qy	41 CysTyRGlyTyRPhaArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer	60
Db	231 TGCTATGTTACTTCAGGAAGCTGAGGAACCTGGTCTGATGCCGAGCTCGAGTGTCA	290
Qy	61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla	80
Db	291 TACGGAAACGGAGGCCACCTGGCATCTATCTCGAGTTTAAAGGAAGCCAGCACCATAGCA	350
Qy	81 GluTyRILESerGlyTyRGlnArgSerGlnProIleTTPilleGlyLeuHisaspProGln	100
Db	351 GAGTACATTAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCTGCACACCACAG	410
Qy	101 LysArgGlnGlnTrpGlnTTPilleAspGlyAlaMetTyRLeuTyRArgSerTrpSerGly	120
Db	411 AAGAGGCAGCAGTGGCAGTGAGTTGATGGGGCCATGTATCTGTACAGATCTCTGGTCTG	470
Qy	121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPhaLeuThr	140
Db	471 AAGTCCATGGTGGGAAACAAGCAGCTGTGCTGAGATGAGCTCCAAATAACAACATTTT	530
Qy	141 TrpSerSerAsnGluCysAsnLysArgGlnHisPhaLeuCysLysTyRArgPro	158
Db	531 TGGAGCAGCAACGATGATGCAACAGCCCACTCTCTGTGCAAGTACCGACCA	584
RESULT 9		
AR100599		
LOCUS	AR100599	1114 bp DNA linear PAT 14-FEB-2000
DEFINITION	Sequence 1 from patent US 6080722.	
ACCESSION	AR100599	
VERSION	AR100599.1	GI:12811047
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	1 (bases 1 to 1114)	
AUTHORS	Soppet,D.R., Li,Y. and Dillon,P.J.	
TITLE	Colon specific gene and protein	
JOURNAL	Patent: US 6080722-A 1 27-JUN-2000;	
FEATURES	Location/Qualifiers	
source	1..1114 /organism="unknown"	
BASE COUNT	288 a 291 c 262 g 273 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	9.01e-83	Length: 1114
Score:	878.00	Matches: 158
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0

```

Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-525-041-2 (1-158) x AB100599 (1-1114)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db 111 ATGGCTTCCAGAACGATCGGCTGCTCTCTATTGCTGAGCTGCCTGCGCAAAACAGGAGTC 170

Qy 21 LeuGlyAspIlelleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
Db 171 CTGGGTGATATCATCATGAGACCCAGCTGCTGCTCTGGATGGTTTACCACCAAGTCCAAT 230

Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db 231 TGTATGGTTACTTACGAAAGCTGAGAACTGGTCTGTATGCCGAGCTCGAGTGTCACTCT 290

Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 291 TACGGAAACGGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 350

Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 351 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTTGGCTTGACCGACCCACAG 410

Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db 411 AAGAGGACGAGTGGCAGTGGATTTGATGGGGCCATGTATCTGTACAGATCTCTGGTCTGGC 470

Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
Db 471 AAGTCCATGGTGGGAAACAGCACTGTGTGAGATGAGCTTCTGTGCAAGTACCGACCA 530

Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db 531 TGGAGCAGCAACGAATGCAACAGCGCAACACTTCTCTGTGCAAGTACCGACCA 584

RESULT 10
AX193504 1114 bp DNA linear PAT 15-AUG-2001
LOCUS
DEFINITION Sequence 1071 from Patent WO0149716.
ACCESSION AX193504
VERSION AX193504.1 GI:15211444
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,
Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
JOURNAL Patent: WO 0149716-A 1071 12-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source
1..1114
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 288 a 291 c 262 g 273 t
ORIGIN
Alignment Scores: 9.01e-83 Length: 1114
Pred. No.: 878.00 Matches: 158
Score: 100.00% Conservativeness: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00%
DB: 6 Gaps: 0

US-09-525-041-2 (1-158) x AX193504 (1-1114)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db 111 ATGGCTTCCAGAACGATCGGCTGCTCTCTATTGCTGAGCTGCCTGCGCAAAACAGGAGTC 170

Qy 21 LeuGlyAspIlelleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
Db 171 CTGGGTGATATCATCATGAGACCCAGCTGCTGCTCTGGATGGTTTACCACCAAGTCCAAT 230

Qy 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db 231 TGTATGGTTACTTACGAAAGCTGAGAACTGGTCTGTATGCCGAGCTCGAGTGTCACTCT 290

Qy 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
Db 291 TACGGAAACGGAGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 350

Qy 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
Db 351 GAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTTGGCTTGACCGACCCACAG 410

Qy 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
Db 411 AAGAGGACGAGTGGCAGTGGATTTGATGGGGCCATGTATCTGTACAGATCTCTGGTCTGGC 470

Qy 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
Db 471 AAGTCCATGGTGGGAAACAGCACTGTGTGAGATGAGCTTCTGTGCAAGTACCGACCA 530

Qy 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
Db 531 TGGAGCAGCAACGAATGCAACAGCGCAACACTTCTCTGTGCAAGTACCGACCA 584

RESULT 11
AX193507 1114 bp DNA linear PAT 15-AUG-2001
LOCUS
DEFINITION Sequence 1074 from Patent WO0149716.
ACCESSION AX193507
VERSION AX193507.1 GI:15211447
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,
Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
JOURNAL Patent: WO 0149716-A 1074 12-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source
1..1114
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 288 a 291 c 262 g 273 t
ORIGIN
Alignment Scores: 9.01e-83 Length: 1114
Pred. No.: 878.00 Matches: 158
Score: 100.00% Conservativeness: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00%
DB: 6 Gaps: 0

US-09-525-041-2 (1-158) x AX193507 (1-1114)

Qy 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
Db 111 ATGGCTTCCAGAACGATCGGCTGCTCTCTATTGCTGAGCTGCCTGCGCAAAACAGGAGTC 170

Qy 21 LeuGlyAspIlelleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40

```

```

|||||
171 CTGGGTGATATCATCATGACCCAGCTGCTGCTCCGGATGGTGTATACCAAGTCCAAT 230
QY
41 CysTyAGlyTyRPhaArgLyLeuAsgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db
231 TGTATGGTTACTTTCAGGAAGCTGAGGAAGTGGTCTGATGCCAGCTCGAGTGTGAGTCT 290
QY
61 TyAGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLeuGluAlaSerThrIleAla 80
Db
291 TACGGAACCGAGGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 350
QY
81 GluTyRlleSerGlyTyRGlNArgSerGlnPrIleTrpIleGlyLeuHisAspProGln 100
Db
351 GAGTACATAGTGGCTATCAGAGAAGCCAGCCCATATGGATTGGCTGTCAGCCACCACAG 410
QY
101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyRLeuTyRArgSerTrpSerGly 120
Db
411 AAGAGGCAGCAGTGGCAGTGGATTGATGGGCCATGATCTGTACAGATCTCTGGTCTGGC 470
QY
121 LysSerMetGlyGlyAsnLyshisCysAlaGluMetSerSerAsnAsnPhelLeuThr 140
Db
471 AAGTCCATGGTGGGAACAAGCAAGCACTGTGCTGAGATGAGCTCCAATAACAACCTTTTAACT 530
QY
141 TrpSerSerAsnGluCysAsnLyshisArgGlnHisPhelLeuCysLyshisTrpArgPro 158
Db
531 TGGAGCAGCAACGAATGCAACAAGCCCAACTTCTGTGCAAGTACCGACCA 584

RESULT 12
AX676920
LOCUS AX676920 1130 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 113 from Patent WO02103028.
ACCESSION AX676920
VERSION AX676920.1 GI:29334486
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Baranova,A.V., Yankovsky,N.K., Kozlov,A.P., Lobashev,A.V. and
Krukovskaya,L.L.
TITLE In silico screening for phenotype-associated expressed sequences
JOURNAL Patented: WO 02103028-A 113 27-DEC-2002;
Biomedical Center (RU)
FEATURES
source
Location/Qualifiers
1..1130
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 291 a 293 c 265 g 281 t
ORIGIN
Alignment Scores:
Pred. No.: 9.16e-83 Length: 1130
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-525-041-2 (1-158) x AX676920 (1-1130)
QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuLeuSerCysAlaLysThrGlyVal 20
Db
123 ATGGCTTCAGAGCATGGGCTGCTCTTATGCTGAGTGGCTGGCCAAACAGGAGTTC 182
QY
21 LeuGlyAspIleLeuMetArgProSerCysAlaProGlyTrpPheTyRHisLysSerAsn 40
Db
183 CTGGGTGATATCATCATGAGCCAGCTGTGCTCTGGATGGTTTTTACCAAGTCCAAT 242
QY
41 CysTyAGlyTyRPhaArgLyLeuAsgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
Db
243 TGTATGGTTACTTTCAGGAAGCTGAGGAAGTGGTCTGATGCCAGCTCGAGTGTGAGTCT 302

```

```

QY 61 TyAGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLeuGluAlaSerThrIleAla 80
Db 303 TACGGAACCGAGGCCACCTGGCATCTATCTGAGTTTAAAGGAAGCCAGCACCATAGCA 362
QY 81 GluTyRlleSerGlyTyRGlNArgSerGlnPrIleTrpIleGlyLeuHisAspProGln 100
Db 363 GAGTACATAGTGGCTATCAGAGAAGCCAGCCCATATGGATTGGCTGTCAGCCACCACAG 422
QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyRLeuTyRArgSerTrpSerGly 120
Db 423 AAGAGGCAGCAGTGGCAGTGGATTGATGGGCCATGATCTGTACAGATCTCTGGTCTGGC 482
QY 121 LysSerMetGlyGlyAsnLyshisCysAlaGluMetSerSerAsnAsnPhelLeuThr 140
Db 483 AAGTCCATGGTGGGAACAAGCAAGCACTGTGCTGAGATGAGCTCCAATAACAACCTTTTAACT 542
QY 141 TrpSerSerAsnGluCysAsnLyshisArgGlnHisPhelLeuCysLyshisTrpArgPro 158
Db 543 TGGAGCAGCAACGAATGCAACAAGCCCAACTTCTGTGCAAGTACCGACCA 596

RESULT 13
AY007243
LOCUS AY007243 1130 bp mRNA linear PRI 23-APR-2001
DEFINITION Homo sapiens regenerating gene type IV mRNA, complete cds.
ACCESSION AY007243
VERSION AY007243.1 GI:12621025
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1130)
AUTHORS Hartupsee,J.C., Zhang,H., Bonaldo,M.F., Soares,M.B. and
Dieckgraefe,B.K.
TITLE Isolation and characterization of a cDNA encoding a novel member of
the human regenerating protein family: Reg IV
JOURNAL Biochim. Biophys. Acta 1518 (3), 287-293 (2001)
MEDLINE 21210973
PUBMED 11311942
REFERENCE 2 (bases 1 to 1130)
AUTHORS Dieckgraefe,B.K., Hartupsee,J.C., Zhang,H., Soares,M.B. and
Bonaldo,M.F.
TITLE Direct Submission
JOURNAL Submitted (18-AUG-2000) Gastroenterology, Washington University
School of Medicine, 660 S. Euclid Ave., Campus Box 8124, St. Louis,
MO 63110, USA
FEATURES
source
Location/Qualifiers
1..1130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="1"
123..599
/note="RegIV"
/codon_start=1
/product="regenerating gene type IV"
/protein_id="AAG02562.1"
/db_xref="GI:12621026"
/translation="MASHSMRLLLLSCLAKTVGLDIIIMPSCAPGFYHKSNCYGY
FKRLRNSDALEQCSYNGAHLASILSKASTAEIYISGYRSQPIWGLHDPQR
QQQWIDGAMTYLRWSGKSMGNKHAEMSSNNFLTWSSNECNKRQHFLEKVRP"
BASE COUNT 291 a 293 c 265 g 281 t
ORIGIN
Alignment Scores:
Pred. No.: 9.16e-83 Length: 1130
Score: 878.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

```

US-09-525-041-2 (1-158) x AY007243 (1-1130).

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 DB 123 ATGGCTTCCAGAGCATGCGGCTGCTCTATTGCTGAGCTGCCCTGGCCAAAACAGGAGTC 182
 QY 21 LeuGlyAspIlelleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 DB 183 CTGGGTGATATCATCATGAGACCCAGCTGTCTCTCGATGGTGTATACCAACAAGTCCAAT 242
 QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 DB 243 TGCATATGTTTCTCAGAAAGCTGAGAACTGGTCTGATGCCGAGCTCGAGTGTCACTCT 302
 QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 DB 303 TACGGAAACGGAGCCACCTGGCATCTATCTCTGAGTTTAAAGGAAGCAGCACCATAGCA 362
 QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 DB 363 GAGTACATAAGTGGCTATCAGAAAGCCAGCCGATATGGATTGGCTGCACGCCACAG 422
 QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 DB 423 AAGAGGCAGCAGTGGCAGTGGATTGATGGGGCCATGTATCTGTACAGATCCTGGTCTGGC 482
 QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
 DB 483 AAGTCCATGGGTGGGAACAAGCATGTGCTGAGATGAGCTCCAATAACAATTTTAACT 542
 QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
 DB 543 TGGAGCAGCAACGAATGCAACAAGCCCAACACTTCTGTGCAAGTACCGACCA 596

RESULT 14
 LOCUS AX193505 1152 bp DNA linear PAT 15-AUG-2001
 DEFINITION Sequence 1072 from Patent WO0149716.
 ACCESSION AX193505
 VERSION AX193505.1 GI:15211445
 KEYWORDS Homo sapiens (human)
 SOURCE Human sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,
 Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
 TITLE Compounds for immunotherapy and diagnosis of colon cancer and
 methods for their use
 JOURNAL Patent: WO 0149716-A 1072 12-JUL-2001;
 CORIXA CORPORATION (US)

FEATURES
 source
 1. .1152
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 296 a 298 c 275 g 283 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 9.35e-83 Length: 1152
 Score: 878.00 Matches: 158
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-525-041-2 (1-158) x AX193505 (1-1152)

QY 1 MetAlaSerArgSerMetArgLeuLeuLeuLeuSerCysLeuAlaLysThrGlyVal 20
 DB 147 ATGGCTTCCAGAGCATGCGGCTGCTCTATTGCTGAGCTGCCCTGGCCAAAACAGGAGTC 206

QY 21 LeuGlyAspIlelleMetArgProSerCysAlaProGlyTrpPheTyrHisLysSerAsn 40
 DB 207 CTGGGTGATATCATCATGAGACCCAGCTGTCTCTCGATGGTGTATACCAACAAGTCCAAT 266
 QY 41 CysTyrGlyTyrPheArgLysLeuArgAsnTrpSerAspAlaGluLeuGluCysGlnSer 60
 DB 267 TGCATATGTTTCTCAGAAAGCTGAGAACTGGTCTGATGCCGAGCTCGAGTGTCACTCT 326
 QY 61 TyrGlyAsnGlyAlaHisLeuAlaSerIleLeuSerLeuLysGluAlaSerThrIleAla 80
 DB 327 TACGGAAACGGAGCCACCTGGCATCTATCTCTGAGTTTAAAGGAAGCAGCACCATAGCA 386
 QY 81 GluTyrIleSerGlyTyrGlnArgSerGlnProIleTrpIleGlyLeuHisAspProGln 100
 DB 387 GAGTACATAAGTGGCTATCAGAAAGCCAGCCGATATGGATTGGCTGCACGCCACAG 446
 QY 101 LysArgGlnGlnTrpGlnTrpIleAspGlyAlaMetTyrLeuTyrArgSerTrpSerGly 120
 DB 447 AAGAGGCAGCAGTGGCAGTGGATTGATGGGGCCATGTATCTGTACAGATCCTGGTCTGGC 506
 QY 121 LysSerMetGlyGlyAsnLysHisCysAlaGluMetSerSerAsnAsnAsnPheLeuThr 140
 DB 507 AAGTCCATGGGTGGGAACAAGCATGTGCTGAGATGAGCTCCAATAACAATTTTAACT 566
 QY 141 TrpSerSerAsnGluCysAsnLysArgGlnHisPheLeuCysLysTyrArgPro 158
 DB 567 TGGAGCAGCAACGAATGCAACAAGCCCAACACTTCTGTGCAAGTACCGACCA 620

RESULT 15
 LOCUS BD062775 1152 bp DNA linear PAT 27-AUG-2002
 DEFINITION Human proteins having secretory signal sequences and DNAs encoding
 these proteins.
 ACCESSION BD062775
 VERSION BD062775.1 GI:22608378
 KEYWORDS JP 2001506484-A/12.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1152)
 AUTHORS Kato,S., Sekine,S., Kimura,T. and Kobayashi,M.
 TITLE Human proteins having secretory signal sequences and DNAs encoding
 these proteins
 JOURNAL Patent: JP 2001506484-A 12 22-MAY-2001;
 SAGAMI CHEMICAL RESEARCH CENTER, PROTEGENE INC
 COMMENT OS Homo sapiens (human)
 PN JP 2001506484-A/12
 PD 22-MAY-2001
 PF 12-SEP-1997 JP 1998513509
 PR 13-SEP-1996 JP 8/243060
 PI SEISHI KATO,SHINGO SEKINE,TOMOKO KIMURA,MIDORI KOBAYASHI PC
 C12N15/12,C07K14/47,C12N15/62
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key
 FT CDS Location/Qualifiers
 1. .1152
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 296 a 298 c 275 g 283 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 9.35e-83 Length: 1152
 Score: 878.00 Matches: 158
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 31, 2003, 09:15:03 ; Search time 72 Seconds
(without alignments)
440.296 Million cell updates/sec

Title: US-09-525-041-2

Perfect score: 878

Sequence: 1 MASRMRLLLLLSCLAKTV.....LTWSSNECNKQHFICKYRP 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 733937 segs, 20641311 residues

Total number of hits satisfying chosen parameters: 733937

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	878	100.0	158	9	US-09-922-217-1070
2	878	100.0	158	9	US-09-922-217-1077
3	878	100.0	158	9	US-09-922-217-1078
4	878	100.0	158	9	US-09-922-217-1079
5	878	100.0	158	9	US-09-922-217-1080
6	878	100.0	158	10	US-09-833-263-1070
7	878	100.0	158	10	US-09-833-263-1077
8	878	100.0	158	10	US-09-833-263-1078
9	878	100.0	158	10	US-09-833-263-1079
10	878	100.0	158	10	US-09-833-263-1080
11	878	100.0	158	12	US-09-525-041-2
12	878	100.0	158	12	US-10-295-027-138
13	878	100.0	158	12	US-10-295-027-781
14	878	100.0	158	12	US-10-295-027-861
15	878	100.0	158	12	US-10-295-027-1200

16	878	100.0	158	14	US-10-025-380-1070
17	878	100.0	158	14	US-10-025-380-1077
18	878	100.0	158	14	US-10-025-380-1078
19	878	100.0	158	14	US-10-025-380-1079
20	878	100.0	158	14	US-10-025-380-1080
21	878	100.0	158	15	US-10-100-608B-2
22	878	100.0	158	15	US-10-157-031-114
23	878	100.0	158	15	US-10-205-823-347
24	878	100.0	166	15	US-10-106-698-6394
25	240	27.3	174	9	US-09-925-297-753
26	231	26.3	174	12	US-10-028-248A-113
27	231	26.3	175	12	US-10-028-248A-110
28	225.5	25.7	166	12	US-09-997-003-30
29	225.5	25.7	166	12	US-09-997-003-43
30	225.5	25.7	174	9	US-09-925-301-1182
31	223	25.4	175	12	US-10-028-248A-38
32	221	25.2	175	12	US-10-028-248A-109
33	221	25.2	175	15	US-10-316-761-3
34	221	25.2	183	9	US-09-925-301-1013
35	218	24.8	174	15	US-10-316-761-4
36	217.5	24.8	158	10	US-09-929-230-11
37	217.5	24.8	158	12	US-10-226-420-11
38	205.5	23.4	151	10	US-09-929-230-2
39	205.5	23.4	151	12	US-10-226-420-2
40	204.5	23.3	152	10	US-09-929-230-5
41	204.5	23.3	152	12	US-10-226-420-5
42	204	23.2	175	10	US-09-978-295A-452
43	204	23.2	175	10	US-09-978-697-452
44	204	23.2	175	10	US-09-978-192A-452
45	204	23.2	175	10	US-09-999-832A-452

ALIGNMENTS

RESULT 1

US-09-922-217-1070
; Sequence 1070, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1070
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1070

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRMRLLLLLSCLAKTVGLDIIIMRSCAPGWFWYHKSNVCYGRKLRNWSDALEQCS 60
DB 1 MASRMRLLLLLSCLAKTVGLDIIIMRSCAPGWFWYHKSNVCYGRKLRNWSDALEQCS 60

QY 61 YNGAHLASILSLKEASTIAEYISGYQORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
Db 61 YNGAHLASILSLKEASTIAEYISGYQORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
QY 121 KSMGNGKHCAMSSNNFLTWSNECNKROHFLCKYRP 158
Db 121 KSMGNGKHCAMSSNNFLTWSNECNKROHFLCKYRP 158

RESULT 2

US-09-922-217-1077
; Sequence 1077, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1077
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1077

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDAELECQS 60
Db 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDAELECQS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYQORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
Db 61 YNGAHLASILSLKEASTIAEYISGYQORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
QY 121 KSMGNGKHCAMSSNNFLTWSNECNKROHFLCKYRP 158
Db 121 KSMGNGKHCAMSSNNFLTWSNECNKROHFLCKYRP 158

RESULT 3

US-09-922-217-1078
; Sequence 1078, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1078
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1078

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDAELECQS 60
Db 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDAELECQS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYQORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
Db 61 YNGAHLASILSLKEASTIAEYISGYQORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
QY 121 KSMGNGKHCAMSSNNFLTWSNECNKROHFLCKYRP 158
Db 121 KSMGNGKHCAMSSNNFLTWSNECNKROHFLCKYRP 158

RESULT 4

US-09-922-217-1079
; Sequence 1079, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1079
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1079

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDAELECQS 60
Db 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNICYGFRKLRNWSDAELECQS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYQORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
Db 61 YNGAHLASILSLKEASTIAEYISGYQORSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
QY 121 KSMGNGKHCAMSSNNFLTWSNECNKROHFLCKYRP 158

Db 121 KSMGKHKCAEMSSNNFLTWSNECKRQHFLCKYRP 158
|||||

RESULT 5

US-09-922-217-1080
; Sequence 1080, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1080
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1080

Query Match 100.0%; Score 878; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCGYFRKLNRWSDAELECS 60
|||
DB 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCGYFRKLNRWSDAELECS 60
|||
QY 61 YNGAHLASILSLKEASTIAEYISGYQRSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
|||
DB 61 YNGAHLASILSLKEASTIAEYISGYQRSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
|||
QY 121 KSMGKHKCAEMSSNNFLTWSNECKRQHFLCKYRP 158
|||
DB 121 KSMGKHKCAEMSSNNFLTWSNECKRQHFLCKYRP 158
|||

RESULT 6

US-09-833-263-1070
; Sequence 1070, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1070
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1070

Query Match 100.0%; Score 878; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCGYFRKLNRWSDAELECS 60
|||
DB 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCGYFRKLNRWSDAELECS 60
|||
QY 61 YNGAHLASILSLKEASTIAEYISGYQRSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
|||
DB 61 YNGAHLASILSLKEASTIAEYISGYQRSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
|||
QY 121 KSMGKHKCAEMSSNNFLTWSNECKRQHFLCKYRP 158
|||
DB 121 KSMGKHKCAEMSSNNFLTWSNECKRQHFLCKYRP 158
|||

RESULT 7

US-09-833-263-1077
; Sequence 1077, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1077
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1077

Query Match 100.0%; Score 878; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCGYFRKLNRWSDAELECS 60
|||
DB 1 MASRMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCGYFRKLNRWSDAELECS 60
|||
QY 61 YNGAHLASILSLKEASTIAEYISGYQRSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
|||
DB 61 YNGAHLASILSLKEASTIAEYISGYQRSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
|||
QY 121 KSMGKHKCAEMSSNNFLTWSNECKRQHFLCKYRP 158
|||
DB 121 KSMGKHKCAEMSSNNFLTWSNECKRQHFLCKYRP 158
|||

RESULT 8

US-09-833-263-1078
; Sequence 1078, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1078

```
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1078

Query Match      100.0%; Score 878; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWYHKSNCGYFRKLRNWSDAELECQS 60
   |||||
Db 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWYHKSNCGYFRKLRNWSDAELECQS 60
   |||||

QY 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
   |||||
Db 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
   |||||

QY 121 KSMGKNKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158
   |||||
Db 121 KSMGKNKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158
   |||||

RESULT 9
US-09-833-263-1079
; Sequence 1079, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833.263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1079
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1079

Query Match      100.0%; Score 878; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWYHKSNCGYFRKLRNWSDAELECQS 60
   |||||
Db 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWYHKSNCGYFRKLRNWSDAELECQS 60
   |||||

QY 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
   |||||
Db 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
   |||||

QY 121 KSMGKNKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158
   |||||
Db 121 KSMGKNKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158
   |||||

RESULT 10
US-09-833-263-1080
; Sequence 1080, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
```

```
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1080
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1080

Query Match      100.0%; Score 878; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWYHKSNCGYFRKLRNWSDAELECQS 60
   |||||
Db 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWYHKSNCGYFRKLRNWSDAELECQS 60
   |||||

QY 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
   |||||
Db 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
   |||||

QY 121 KSMGKNKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158
   |||||
Db 121 KSMGKNKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158
   |||||

RESULT 11
US-09-525-041-2
; Sequence 2, Application US/09525041
; Publication No. US20030158098A1
; GENERAL INFORMATION:
; APPLICANT: Soppet et al.
; TITLE OF INVENTION: Colon Specific Gene and Protein
; FILE REFERENCE: PFI78D2
; CURRENT APPLICATION NUMBER: US/09/525,041
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 09/162,508
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: US 08/468,413
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-525-041-2

Query Match      100.0%; Score 878; DB 12; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWYHKSNCGYFRKLRNWSDAELECQS 60
   |||||
Db 1 MASRSMRLLLLLCLAKTGVLDIIMRPSCAPGWYHKSNCGYFRKLRNWSDAELECQS 60
   |||||

QY 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
   |||||
Db 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
   |||||

QY 121 KSMGKNKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158
   |||||
Db 121 KSMGKNKHCAMSSNNFLTWSSNECNKQHFCLKYRP 158
   |||||

RESULT 12
US-10-295-027-138
; Sequence 138, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
```

APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 138
LENGTH: 158
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-781

Query Match 100.0%; Score 878; DB 12; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRSCAPGWFYHKSNCYGFYFKLRNWSDAELECS 60
DB 1 MASRSMRLLLLSCLAKTGVLDIIMRSCAPGWFYHKSNCYGFYFKLRNWSDAELECS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPQKQOWIDGAMLYRSWSG 120
DB 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPQKQOWIDGAMLYRSWSG 120
QY 121 KSMGKNKHCAMSSNNFLTWSNECNKQHFCLKYRP 158
DB 121 KSMGKNKHCAMSSNNFLTWSNECNKQHFCLKYRP 158

RESULT 13
US-10-295-027-781
Sequence 781, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.

APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 781
LENGTH: 158
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-781

Query Match 100.0%; Score 878; DB 12; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASRSMRLLLLSCLAKTGVLDIIMRSCAPGWFYHKSNCYGFYFKLRNWSDAELECS 60
DB 1 MASRSMRLLLLSCLAKTGVLDIIMRSCAPGWFYHKSNCYGFYFKLRNWSDAELECS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPQKQOWIDGAMLYRSWSG 120
DB 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPQKQOWIDGAMLYRSWSG 120
QY 121 KSMGKNKHCAMSSNNFLTWSNECNKQHFCLKYRP 158
DB 121 KSMGKNKHCAMSSNNFLTWSNECNKQHFCLKYRP 158

RESULT 14
US-10-295-027-861
Sequence 861, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733

;; PRIOR FILING DATE: 2000-09-15
;; PRIOR APPLICATION NUMBER: US 60/350,666
;; PRIOR FILING DATE: 2001-11-13
;; PRIOR APPLICATION NUMBER: US 60/335,394
;; PRIOR FILING DATE: 2001-11-15
;; PRIOR APPLICATION NUMBER: US 60/332,464
;; PRIOR FILING DATE: 2001-11-21
;; PRIOR APPLICATION NUMBER: US 60/334,393
;; PRIOR FILING DATE: 2001-11-29
;; PRIOR APPLICATION NUMBER: US 60/340,376
;; PRIOR FILING DATE: 2001-12-14
;; PRIOR APPLICATION NUMBER: US 60/347,211
;; PRIOR FILING DATE: 2002-01-08
;; PRIOR APPLICATION NUMBER: US 60/356,714
;; PRIOR FILING DATE: 2002-01-08
;; PRIOR APPLICATION NUMBER: US 60/347,349
;; PRIOR FILING DATE: 2002-01-10
;; PRIOR APPLICATION NUMBER: US 60/355,250
;; PRIOR FILING DATE: 2002-02-13
;; PRIOR APPLICATION NUMBER: US 60/356,714
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 1386
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 861
;; LENGTH: 158
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-295-027-861

Query Match 100.0%; Score 878; DB 12; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASRSMLLLLSCLAKTGVLDGIIMRPSCAPGWYHKSNICYGFRKLRNWSDALEECQS 60
DB 1 MASRSMLLLLSCLAKTGVLDGIIMRPSCAPGWYHKSNICYGFRKLRNWSDALEECQS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
DB 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
QY 121 KSMGKNKHCAMSSNNFLTWSSNECNKQHFCLCKYRP 158
DB 121 KSMGKNKHCAMSSNNFLTWSSNECNKQHFCLCKYRP 158

RESULT 15

US-10-295-027-1200
;; Sequence 1200, Application US/10295027
;; Publication No. US20030232350A1
;; GENERAL INFORMATION:
;; APPLICANT: Afar, Daniel
;; APPLICANT: Aziz, Natasha
;; APPLICANT: Gineberg, Wendy M.
;; APPLICANT: Gish, Kurt C.
;; APPLICANT: Glynn, Richard
;; APPLICANT: Hevezi, Peter A.
;; APPLICANT: Mack, David H.
;; APPLICANT: Murray, Richard
;; APPLICANT: Watson, Susan R.
;; APPLICANT: Eos Biotechnology, Inc.
;; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
;; FILE REFERENCE: 018501-012500US
;; CURRENT APPLICATION NUMBER: US/10/295,027
;; CURRENT FILING DATE: 2002-11-13
;; PRIOR APPLICATION NUMBER: US 09/663,733
;; PRIOR FILING DATE: 2000-09-15
;; PRIOR APPLICATION NUMBER: US 60/350,666
;; PRIOR FILING DATE: 2001-11-13
;; PRIOR APPLICATION NUMBER: US 60/335,394
;; PRIOR FILING DATE: 2001-11-15
;; PRIOR APPLICATION NUMBER: US 60/332,464
;; PRIOR FILING DATE: 2001-11-21

;; PRIOR APPLICATION NUMBER: US 60/334,393
;; PRIOR FILING DATE: 2001-11-29
;; PRIOR APPLICATION NUMBER: US 60/340,376
;; PRIOR FILING DATE: 2001-12-14
;; PRIOR APPLICATION NUMBER: US 60/347,211
;; PRIOR FILING DATE: 2002-01-08
;; PRIOR APPLICATION NUMBER: US 60/347,349
;; PRIOR FILING DATE: 2002-01-10
;; PRIOR APPLICATION NUMBER: US 60/355,250
;; PRIOR FILING DATE: 2002-02-08
;; PRIOR APPLICATION NUMBER: US 60/356,714
;; PRIOR FILING DATE: 2002-02-13
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 1386
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 1200
;; LENGTH: 158
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-295-027-1200

Query Match 100.0%; Score 878; DB 12; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.6e-86;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASRSMLLLLSCLAKTGVLDGIIMRPSCAPGWYHKSNICYGFRKLRNWSDALEECQS 60
DB 1 MASRSMLLLLSCLAKTGVLDGIIMRPSCAPGWYHKSNICYGFRKLRNWSDALEECQS 60
QY 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
DB 61 YNGAHLASILSLKEASTIAEYISGYQSQPIWIGLHDPKQKQOWIDGAMLYRSWSG 120
QY 121 KSMGKNKHCAMSSNNFLTWSSNECNKQHFCLCKYRP 158
DB 121 KSMGKNKHCAMSSNNFLTWSSNECNKQHFCLCKYRP 158

Search completed: December 31, 2003, 09:24:39
Job time : 73 secs